

Detailed analysis of public RNAseq data and long non-coding RNA: a proposed enhancement to mesenchymal stem cell characterisation

Sebastien Riquier¹, Marc Mathieu¹, Anthony Boureux¹, Florence Ruffle¹, Jean-Marc Lemaitre¹, Farida Djouad¹, Nicolas Gilbert¹ and Therese Commes^{1,*}

February 27, 2020

Abstract

The development of RNA sequencing (RNAseq) and corresponding emergence of public datasets have created new avenues of transcriptional marker search. The long non-coding RNAs (lncRNAs) constitute an emerging class of transcripts with a potential for high tissue specificity and function. Using a dedicated bioinformatics pipeline, we propose to construct a cell-specific catalogue of unannotated lncRNAs and to identify the strongest cell markers. This pipeline uses *ab initio* transcript identification, pseudoalignment and new methodologies such as a specific k-mer approach for naive quantification of expression in numerous RNAseq data.

For an application model, we focused on Mesenchymal Stem Cells (MSCs), a type of adult multipotent stem-cells of diverse tissue origins. Frequently used in clinics, these cells lack extensive characterisation. Our pipeline was able to highlight different lncRNAs with high specificity for MSCs. *In silico* methodologies for functional prediction demonstrated that each candidate represents one specific state of MSCs biology. Together, these results suggest an approach that can be employed to harness lncRNA as cell marker, showing different candidates as potential actors in MSCs biology, while suggesting promising directions for future experimental investigations.

1 Introduction

1

The increasing popularity of RNAseq and the ensuing aggregation of this data-type into public databases enable the search for new biomarkers across large cohorts of donors or cell types for the purpose pathological conditions or cellular lineages identification. As such, RNAseq has paved the way for the discovery of novel transcriptional biomarkers such as long noncoding RNAs (lncRNAs), that have emerged as a fundamental molecular class. A growing number of lncRNAs have been identified in the last decades, with their number approaching that of coding RNAs (17910 annotated human lncRNAs in the latest v32 version of the GENCODE versus 19 965 coding genes).

*To whom correspondence should be addressed: Tel: +33 046 7335711 ; Email: therese.commes@inserm.fr

¹IRMB, University of Montpellier, INSERM, Montpellier, France.

An increasing body of evidence has highlighted characteristics that define lncRNAs as therapeutic targets as well as potential tissue-specific markers [1]. Indeed, despite their non-coding nature, a large spectrum of functional mechanisms have been associated to lncRNAs [2, 3]. These include: endogenous competition (miRNA sponging for example), protein complex scaffolding and guides for active proteins with RNA-DNA homology interactions. These mechanisms occur in various physiological or pathological processes such as development, cancer and immunity [4, 5, 6]. To date, there is no finite list of long non-coding isoforms, making it difficult to construct a complete lncRNA catalogue due to the high number of transcripts and their tissue-specific expression [7, 8]. The absence of a complete catalogue makes it difficult to establish a comprehensive lncRNA expression profile. Thus, currently, the best strategy for the study of lncRNAs consists in the prediction of transcripts from a selection of RNAseq data in a tissue-specific condition. This strategy was successful in novel lncRNA biomarker discovery in pathological conditions [9, 10], but was poorly explored for cell lineage characterisation. Taking into account their functional importance and specificity, these RNAs should therefore not be ignored in establishing the molecular identity of a cell type.

Cell characterisation by specific markers bring different challenges such as the importance of probing the specificity of the marker and its limits in an extended number of various cell types, rather than using a control/patient experimental model. Moreover, the cells are not in a fixed state and display a variable transcriptional activity depending on cell status, environment, culture conditions and other parameters [1]. Furthermore, the lncRNAs' function is generally poorly assessed, except in the case of recurrent known transcripts (HOTAIR, H19). Thus, the *in silico* elaboration of a lncRNA catalogue that document the functional domains where the candidates could act, will be beneficial in the identification of lncRNAs' role and thus, in future experiments.

To this end, we have developed an integrated four-step procedure consisting of: i) an *ab initio* transcript reconstruction from RNAseq data and characterisation of novel transcripts, ii) a differential analysis using pseudoalignment coupled with a machine learning solution in order to extract the most cell-specific candidates, iii) an original step of tissue-expression validation with specific k-mers search in large and diversified transcriptomic datasets iv) an in-depth analysis to predict lncRNAs' functional potential from *in silico* prediction approaches. The notable advantage of introducing an *in silico* verification using k-mers is to allow a precise and in-depth determination of lncRNAs expression profile and to quickly interrogate their lineage specificity. In addition to that, validation of newly identified lncRNAs has been undertaken using RT-qPCR and long read sequencing (with Oxford Nanopore technology).

Mesenchymal Stem Cells (MSCs) are defined as multipotent adult stem cells, harvested from various tissues, including Bone Marrow (BM), Umbilical Cord (UC), and Adipose tissue (Ad). MSCs are an interesting cell type to explore since these cells lack the extended transcriptional characterisation that could highlight their lineage belonging and/or the possibility of distinguishing them from other mesodermal cell types such as fibroblasts and pericytes [11, 12]. The commonly admitted surface markers for MSCs, proposed by the International Society for Cellular Therapy (ISCT), and required to identify MSCs since 2006 are THY1 (CD90), NT5E (CD73), ENG (CD105) concerning the positive markers, and CD45, CD34, CD14 or CD11b, CD79alpha or CD19 and HLA-DR concerning the negative markers [13]. These markers are not distinctive and

may therefore not be sufficient for the definition of cellular or biological properties. Considering their different therapeutic properties (chondro and osteo differentiation potential, immunomodulation and production of trophic factors) [14] and given the increasing usage of these cells for academic and preclinical research [15], a detailed molecular characterisation of MSCs and predictive marker of functionality will constitute an important tool in regenerative medicine. lncRNAs have emerged as a class of transcripts with tissue-specific expression and important functions, such as the regulation of MSCs function [16, 17, 18], and remain largely unexplored in these cells.

To address this need, we performed a broad transcriptomic analysis of novel lncRNAs on human Mesenchymal Stem Cells (MSCs). We started from publicly available MSCs RNAseq, selecting ribodepleted datasets in order to enhance lncRNAs discovery and to explore the poly-(A)+ as poly(A)- lncRNAs. We restricted the differential expression analysis to a bone-marrow MSCs source compared to "nonMSC" cell counterparts. Once achieved, in depth *in silico* analysis was performed to check the lncRNA cell specific profiles with more and extensive datasets. To validate our approach, RNAseq data from eight publicly available libraries of normal MSCs containing a large diversity of noncancerous cell types were used for novel lncRNAs detection and tissue expression comparison. We initially reconstructed more than 70000 unannotated lncRNAs present in human bone-marrow MSCs. These lncRNAs were assigned depending on their position relative to annotated genes: MSC-related Long Intergenic Non Coding transcripts named "Mlinc", and MSC-related Long Overlapping Antisense transcripts called "Mloanc". Among them, 35 Mlincs were specifically enriched in the cell lineage compared to the "non-MSC" counterpart group. Finally, after a further selection of the three most specific Mlincs, detailed *in vitro* and *in silico* functional exploration were performed.

2 Material and methods

2.1 Data collection and basic processing

The public RNAseq datasets (in fastq format) have been assessed using the ENCODE, EBI ArrayExpress service or SRA database at each step of the pipeline:

i) lncRNAs prediction and first differential analysis (Table S1), ii) k-mer search in ENCODE data to refine lncRNAs' specificity (Table S2), iii) k-mer search in FANTOM6 CAGE dataset and single cell analysis (scRNAseq) from Adipose MSCs by X. Liu et al raw data [19] for functional investigations (Table S3), iv) k-mer search in MSCs in different conditions (Table S4).

The reads quality were assessed with FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) to avoid the implementation of poor quality data in the analysis. Data from Peffers et al. [20], added to ENCODEs BM-MSCs RNAseq data, were selected for the MSClinc and Mloanc characterisation and the differential step analysis considering the above-mentioned features: Ribo-zero technology, stranded and paired-ends RNAseq. Peffers data had a forward-reverse library orientation instead of a reverse-forward orientation of a classic Illumina sequencing, thereby the order of paired files was manually reversed. The fastq files used for lncRNAs prediction, referred as "MSC" group, were used for the differential analysis against the other cell types as "non-MSC" group, (Table S1) were mapped using CRAC v2.5.0 software [21] on the indexed GRCh38 human genome, including mitochondria, with stranded, -k 22 and rf options.

2.2 *Ab initio* assembly for transcripts prediction or unannotated transcripts prediction

The aligned reads of the "MSC" group were put through *ab initio* transcript assembly. Unannotated transcripts were predicted with the following procedure: i/an *ab initio* reconstruction was performed on individual RNAseq with the StringTie [22] version 1.3.3b, with -c 5 -j 5 rf -f 0.1 (5 spliced reads are necessary to predict a junction and 5 reads at minimum are required to predict an expressed locus), ii/ the output individual gtf files obtained with the RNAseq of "MSC" group were then merged with the StringTie version 1.3.3b with -f 0.01 -m 200 and with a minimum TPM of 0.5, with the Ensembl human annotation (hg38) v90 used as guide for StringTie. The GTF was parsed with BEDTools [23] to dissociate new intergenic lncRNAs (lincRNAs) from annotated RNAs (coding or annotated lncRNAs), by applying filter criteria classically used in lncRNAs prediction [24], excluding transcript models overlapping (by 1 bp or more) any annotated coordinates. The resulting GTF of unannotated lincRNA from MSCs is referred as "Mlinc".

In parallel, the GTF was parsed with BEDTools to dissociate overlapping-antisens lncRNAs (named Mloanc), by applying filter criteria classically used in lncRNAs prediction, keeping transcript overlapping any annotated coordinates, then excluding transcript models overlapping these annotated coordinates on the same strand. The resulting GTF of MSCs overlapping-antisens lncRNAs is referred as "Mloanc" (Figure 1).

2.3 Long-read sequencing

The library was generated with 250 ng polyA+ mRNA purified from 50 μ g of human BM-MSCs total RNA. The polyA+ mRNA were treated according to the cDNA-PCR sequencing kit protocol (ref SQK-PCS108) as recommended by Oxford Nanopore. 3 254 396 sequences were obtained on the Oxford Nanopore Minion sequencer. The base calling was done with albacore version 2.2.7. 2 720 928 long-reads were successfully mapped using Minimap2 [25] version 2.10-r764 on GRCh38 human genome with default options used for Oxford Nanopore technology.

2.4 Quantification with pseudoalignment and feature selection

Kallisto v0.43.1 [26] was used directly on RNAseq raw fastq from the "MSC" and "non-MSC" groups. This pseudoalignment was performed with a number of bootstraps (-b) of 100, using a Kallisto index containing the sequences of all transcripts: the Ensembl coding and non-coding transcripts (v90) plus the predicted lincRNA and lncRNAs. Sleuth version 0.29.0 [27] was used with R for differential expression statistical analysis using the Walt test method, to compare the "MSC" group against the "non-MSC" group (including Lymphocytes, Macrophages, Hepatocytes, IPS, ESCs, HUVECs, Neurons, chondrocytes). Analysis was performed at the gene level for the annotated genes and at the transcript level for the predicted lincRNA and lncRNAs. Genes or lncRNAs having a log2 fold-change between "MSC" and others greater than 0.5 and a p-value lower than or equal to 0.05 were selected. Finally, only transcripts/genes overexpressed in MSCs were selected. Each category (annotated transcripts, lincRNAs and lncRNAs) of potential candidates passing the first differentiation expression filter were separated for feature selection analysis. Boruta 6.0 [28] was used with 10000 maximum runs and a pvalue of 0.01 on each category, with multiple comparisons adjustment using the Bonferroni method (mcAdj = TRUE). Candidates passing the boruta test as "Confirmed" for each category were selected as reliable biomarkers.

2.5 Quantification by k-mers search

To quantify the expression of a transcript or a gene in available RNAseq with a rapid procedure, specific k-mers of 31nt length were extracted from the candidate sequence. A specific k-mer of an annotated candidate corresponds to a 31nt sequence that maps once on the genome and once on the reference transcriptome (Ensembl v90). In case of unannotated transcript (Mlinc, Mloanc) a specific k-mer maps once on the genome and is absent from reference transcriptome. The automated selection of specific k-mers is ensured by the Kmerator tool (in preparation) (<https://github.com/Transipedia/kmerator>). The k-mers were then quantified directly in raw fastq files using countTags (<https://github.com/Transipedia/countTags>). The quantification is expressed by the average count of all k-mers for one transcript, normalised by million of total k-mers in the raw file.

In FANTOM6 Dataset (<https://doi.org/10.1101/700864>, article not peer-reviewed) containing CAGE Analysis, to approach a "Transcript Per Millions" normalisation, the number of k-mers quantified was normalised by the total number of reads in million.

2.6 Genomic intervals assessment

DNase-seq intervals of enrichment were directly downloaded from ENCODE in bed format for BM-mesenchymal cells (ENCFF832FHZ) and hematopoietic progenitors (ENCFF378FCS). The H3K27ac (GSM3564514) and H3K4me3 (GSM3564510) ChIP results from undifferentiated BM-MSCs of the Agrawal Singh S. et al. study [29] were downloaded from GEO database in wig format, and remapped to the hg38 genome with CrossMap (<http://crossmap.sourceforge.net/>).

2.7 *in silico* functional prediction

We used LncADeep [30] to identify particular correlations between candidates and proteins. Beginning with our selection of 3 candidates, we filtered shared predicted proteins and selected protein uniquely predicted as interacting uniquely with the concerned candidate. The pathways concerned with these unique protein were identified with reactome.

Tarpmir was used to identify possible target site of human miRNA from miRbase ($p = 0.5$) [31] and FEELnc [32] to decipher the coding potential of candidates, using the coding and non-coding par of Ensembl annotation sequences as model.

2.8 Single-cell analysis

Single-cell data were pseudoaligned with Kallisto, with the same index used for the initial bulk RNAseq analysis. Pseudoalignment of 10X genomics data, correction, sorting and counting was made by Kallisto "bus" functions. Count matrices were processed with Seurat R package [33, 34]. Empty droplets were estimated by barcode ranking knee and inflection points, only droplet with a minimal count of 10000 were kept. In the end, 26071 droplets remain.

After normalisation, Inter-donor batch effect was corrected with ComBat method in sva R package [35] (Combat function, prior.plots=FALSE, par.prior=TRUE). Cell cycle scoring was made by CellCycleScoring Seurat function, using gene set used by the initial authors [19]. Finally, other sources of unnecessary variability as percent of mitochondrial genes, cell cycle and number of UMIs were regressed using ScaleData Seurat function.

To decipher genes enriched in cells positive for our markers, cells with a scaled expression superior or equal to 0.1 were labelled as positives, whereas cells with an expression inferior to the level were labelled as negatives. Then, markers of these cell were deciphered using FindAllMarkers Seurat function with a minimum FC threshold of 0.15. Expression of our markers in the Ad-MSCs population was made by FeaturePlot Seurat function after UMAP dimensional reduction, the gene enrichments were represented with VlnPlot function.

2.9 Data visualisation

Genome browser-like figures were generated with Gviz R package [36]. Bam tracks were generated from merged BAMs used for transcript prediction. Heatmaps were generated using superHeat R package (<https://github.com/rlbarter/superheat>).

2.10 Ethics approval and consents

Human primary MSCs was obtained from patients who had granted the authors written informed consent with approval of the General Direction for Research and Innovation, the department in responsible for questions of ethics within the French Ministry of Higher Education and Research (registration number: DC-2009-1052). Human primary myoblasts were collected from patients of the CHU of Montpellier, France (the Montpellier University Hospital) who had provided informed consent. All experiments were performed in accordance with the Declaration of Helsinki and approved by the ethical committee of the CHU of Montpellier (France). Samples were approved for storage by the French "Ministre de l'Enseignement et de la Recherche" (NDC-2008-594). Liver samples were obtained from the Biological Resource Center of Montpellier CHU (CRB-CHUM; <http://www.chu-montpellier.fr>; Biobank ID: BB-0033-00031). The procedure was approved by the French Ethics Committee and written or oral consent was obtained from the patients.

2.11 Cell preparation and culture conditions

MSCs were isolated from bone marrow aspirates of patients undergoing hip replacement surgery, as previously described [37]. Cell suspensions were plated in α -MEM supplemented with 10 % FCS, 1 ng/mL FGF2 (R&D Systems), 2 mM L-glutamine, 100 U/mL penicillin and 100 μ g/mL streptomycin. These were shown to be positive for CD44, CD73, CD90 and CD105 and negative for CD14, CD34 and CD45 and used at the third or fourth passage. Human skin fibroblasts were cultured in DMEM high glucose supplemented with 10 % FCS. For Ad-MSCs isolation, adipose tissue was digested with 250 U/mL collagenase type II for 1 h at 37 °C and centrifuged (300 g for 10 min) using routine laboratory practices. The stroma vascular fraction was collected and cells filtered successively through a 100 μ m, 70 μ m and 40 μ m porous membrane (Cell Strainer, BD-Biosciences, Le-Pont-de-Claix, France). Single cells were seeded at the initial density of 4000 cell/cm² in α MEM supplemented with 100 U/mL penicillin/streptomycin (PS), 2 mmol/mL glutamine (Glu) and 10% fetal calf serum (FCS). After 24 h, cultures were washed twice with PBS. After 1 week, cells were trypsinised and expanded at 2000 cells/cm² till day 14 (end of passage 1), where Ad-MSCs preparations were used.

Human umbilical vein endothelial cells (HUVEC) obtained from Clonetics (Lonza, Levallois Perret, France) were cultured in complete EGM-2MV (Lonza) supplemented with 3 % FCS (HyClone; Perbio Science, Brebieres, France)

Primary human myoblasts were isolated and purified from skeletal muscles of donors, as described by Kitzmann et al [38]. Purified myoblasts were plated in Petri dishes and cultured in growth medium containing Dulbecco's Modified Eagle's Medium (Gibco) supplemented with 20 % foetal bovine serum (FBS) (GE Healthcare, PAA), 0.5 % Ultrosor G serum substitute (PALL life sciences) and 50 μ g/ml Gentamicin (Thermo Scientific, France) at 37°C in humidified atmosphere with 5 % CO₂. All experiments were carried out between passage 4 (P4) and P8 to avoid cell senescence.

IPSCs were maintained in mTeSR-1TM medium (STEMCELL Technologies), in Petri dishes with matrigel (Corning, France). For the passages, cells were incubated in Gentle Cell Dissociation Reagent (STEMCELL Technologies) at room temperature, dissociation medium was discarded and cells incubated in mTeSR medium. All cell cultures were performed at 37°C with 5% of O₂ and 10% of CO₂.

Primary human hepatocytes (PHHs) were isolated, as described previously [39], from liver resections performed in adult patients.

NSC derived from H9 or directly bought (StemPro) have been cultivated on laminine with StemPro NSC SFM medium.

H9 embryonic stem cells were cultivated in ESICO medium in a coculture H9/MEF (Mouse Embryonic Fibroblasts) at 37 °C with 5% of O₂ and 5% CO₂.

2.12 RNA preparation and reverse transcription

Total RNA was isolated using TRIzol reagent (invitrogen) or RNeasy Mini Kit (Qiagen, France) according to the manufacturer protocol. RNA was quantified using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, France). RNA quality and quantity were further assessed using the 2100-Bioanalyzer (Agilent Technologies, Waldronn, Germany). Only preparations with RNA integrity number (RIN) values above 7 were considered. Reverse-transcription was performed either with random hexamers using the GeneAmp Gold RNA PCR Core kit (Applied Biosystems) or with oligo(dT) using SuperScript™ First-Strand Synthesis System for RT-qPCR (invitrogen, France).

2.13 Real-time quantitative PCR

Primer pairs were designed with primer3 online software (<http://bioinfo.ut.ee/primer3-0.4.0/>) from the transcripts' sequences. Primer pairs with a perfect and unique match on the human genome were validated with ucsc blat software (<https://genome.ucsc.edu>). As a final verification, primers were visualised in parallel with the bam alignment using IGV (<http://software.broadinstitute.org/software/igv/>) to verify that the primers overlap zones with read coverage. If possible, primer-pairs were designed to span an intron when present in the genomic sequence. Primers were designed for a mean T_m of 60°C. Quantitative PCR (qPCR) were performed using LightCycler 480 SYBR Green I Master mix and real-time PCR instrument (Roche). PCR conditions were 95 °C for 5 min followed by 45 cycles of 15 s at 95 °C, 10 s at 60 °C and 20 s at 72 °C. For each reaction, a single amplicon with the expected melting temperature was obtained.

The gene encoding ribosomal protein S9 (RPS9) was used as house-keeping gene for normalisation. The threshold cycle (C_t) of each amplification curve was calculated by Roche's LightCycler 480 software using the second derivative maximum method. The relative amount of transcripts were calculated using the ddC_t method [40].

Results

For the purpose of generating a catalogue of all transcripts in any particular cell type, we developed a pipeline for the characterisation of all RNAs and their expression profile in a large collection of RNAseq data. The procedure includes four steps: i) an *ab initio* transcripts reconstruction from RNAseq data and identification of unannotated transcripts, ii) a differential analysis using pseudoalignment coupled with a machine learning solution in order to extract the most cell-specific candidates, iii) an original step of tissue- expression validation with a kmer approach (comparing large transcriptomic datasets), iv) an in-depth analysis to predict lncRNAs functional potential from *in silico* prediction approaches (Figure 1). To illustrate the procedure, we produced a RNA catalogue from bone-marrow MSCs ("MSC" group).

2.14 General features of the predicted MSC catalogue of lncRNAs

As mentioned above, we started with the *ab initio* reconstruction of any transcript from bone-marrow RNAseq with Stringtie assembler after mapping the reads with the CRAC software (see Materials and Methods for parameters). New isoforms of annotated transcripts were ignored. Of the 200 243 transcripts present in Ensembl annotation (version 90), 105 511 (52.6%) were detected in MSCs (TPM >0.1 in pseudoalignment quantification).

73 463 new lncRNAs were reconstructed. This fraction of unannotated transcripts represent 41% of detected transcripts, so in our case the *ab initio* reconstruction made it possible to almost double the inventory of detectable signatures in MSCs (Figure 2A). Of these, 34 712 were found to be intergenic and were thus referred to as "Mlinc" RNAs (MSC-related long intergenic non-coding RNAs), and 38 751 were found to be overlapping coding regions but in anti-sense orientation and thus referred to as "Mloanc" RNAs (MSC-related long overlapping antisense non-coding RNAs, with criteria as described in materials and methods and Figure 2A).

The *ab initio* method by itself is not sufficient to efficiently determine the lncRNAs' full length sequences. Moreover, this step does not preclude the possibility of false positives and at this point of the analysis, all the different rebuilt transcripts are considered to be windows of RNA expression or possible artefacts. These candidates are filtered and, for most interesting candidates, their true form is to be refined through experimental methods. We also assessed the general characteristics of predicted *de novo* lncRNAs in MSCs. Globally, Mlincs and Mloancs are shorter transcripts with longer exons compared to coding genes and annotated lncRNAs. The large majority of predicted lncRNAs are mono exonic (99% for Mlincs, 79% for Mloancs), with a length close to 200nt (Figure 2B-C). A consequence of the abundance of mono-exonic lncRNAs is an infinitesimally small number of variant forms. Only 0.15% and 0.82% of Mlincs and Mloancs are not mono isoforms, respectively.

The GC content of reconstructed lncRNAs is lower than that of coding or non-coding annotated genes (Figure 2D). This low GC proportion of around 40% is a common feature in *ab initio* transcript prediction, observed in a majority of studies of different species, from mammals, insects, plants or prokaryotes [41, 42, 43, 44].

2.15 Enrichment of a restricted set of Mlincs and Mloancs

In this second step, our objective was to obtain a restricted set of potential transcripts, using successive filtering approaches that would reveal their cell specificity. We quantified annotated transcripts, Mlincs and Mloancs with kallisto pseudoalignment [26] in a cohort constituted of two groups : the "MSC" group contained the BM-MSCs initially used for *ab initio* reconstruction,

and the "non-MS" group used for comparison, composed of a large panel of different cell types including hESC, hematopoietic precursors and stem cells, primary chondrocytes, IPS, hepatocytes, neurons, lymphocytes and macrophages (Table S1).

Only over-expressed transcripts in "MS" group versus "nonMS" group were selected. Differential statistical tests were made with Sleuth, a tool specially dedicated to Kallisto results [27] (see all selective parameters in Materials and Methods). We performed two differential expression analyses: one at the gene level for Ensembl annotation, and the other at the transcript level for unannotated transcripts, to give the most likely variant form of the predicted lncRNAs. After this differential analysis, 2801 annotated genes, 655 Mlincs and 3032 Mloancs are significantly overexpressed in BM-MSs (Figure 2 E-F).

The lncRNAs are commonly known to be less expressed than coding genes and this was observed in our selected annotated genes and new lncRNAs (Figure 2G). As a validation of our procedure, we found the 3 positive MSs markers of ISCT among the selected annotated genes: THY1 (CD90), ENG (CD105), and NT5E (CD73). We also retrieved some influencers of MSs activity, for example WNT5A [45, 46], Lamin A/C [47], FAP [48] and others. The complete list of selected genes is provided in Table S5.

2.16 Feature selection for the most discriminating coding and non-coding markers

In an attempt to select the best candidates, we retained lncRNAs with the most discriminative profile between "MS" and "non-MS" groups. In our case, the limitation with a classical "top" ranking by Fold Change (FC) or p-value is the presence of subgroups of different types of cells inside the negative "non-MS" group. The FC, estimated by the Beta value in Figure 2C, appears to be a biased indicator of differential expression as it can select strong but localised expressed lncRNAs in cells poorly represented in our negative group, leading to potential false positive results.

To avoid this problem, we used the Boruta feature selection (see material and method section), for selection of discriminating features based on "random forest" machine learning methodology. Boruta was used separately on each group of candidates (annotated genes, Mlincs and Mloancs) to extract a restricted representation of the most relevant MSs signatures. The top 35 importance scores were selected for genes, Mlincs and Mloancs. We arbitrary chose to select the first 35 transcripts for each group based on our observation of a drop in the importance score in coding the gene series. Considering the expression profile of the top 35 coding genes and predicted Mlincs, BM-MSs clusterised independently from other cells types (Figure 3A).

In contrast, the selection of Mloancs did not provide a satisfying clustering as they had similar expression profiles in MSs and other closely related cell types, in particular in primary chondrocytes (Figure S1). For this reason, Mloancs were not retained for further analysis. Selected annotated genes showed a poor specificity, with only few candidates showing a clear difference of expression between MS and others: APCDD1L, HOTAIR, KRTAP1-5 and SMILR. The 3 positive MSs markers from the ISCT were absent in this selection. The novel top 35 Mlincs showed less expression overall but with a more distinctive profile and a higher number of possible MSs markers with clear contrast of expression. The characteristics, genomic intervals and sequences of the 35 candidates are presented in Table S7.

To assess the potential of genes already proposed as potential MSs biomarkers by ISCT (Figure S2) or other potential MSs markers proposed by different authors [14] (Figure S3), we

made a separated expression heatmap, without filter. Among these previously proposed markers, THY1 (CD90) presented the most specific profile. However, each gene presented expression in distinct non-MSCs types.

2.17 Validation of selected Mlincs with long reads sequencing

As mentioned above, classical annotation of lncRNAs with *ab initio* short read methods suffers from inaccuracies and biases. The Oxford Nanopore long read technology (ONT) can sequence entire cDNA, which constitutes a clear technological advantage, not only in confirming the existence of the transcripts but also as it makes it possible to precisely identify the genomic intervals of lncRNA candidates. We performed long-read sequencing of a poly(A) RNA library obtained from a BM-MSCs sample. Among the top 35 selected Mlincs, with around 3 million total reads, 4 transcripts are covered with the ONT sequencing.

These intergenic lncRNAs are named as Stringtie output (SetName. TranscriptNumber. VariantNumber): Mlinc.28428.2, MlincV4.128022.2, MlincV4.89912.1 and MlincV4.64225.1. To support the above transcriptional units, we compared them with our short read data and searched for epigenetic status at the locus of the Mlincs in bone-marrow stromal mesenchymal cells. We looked at DNase sensitive site, H3K27 acetylation, H3K4 trimethylation that commonly corresponds to active regulatory regions (Figure 3B). We globally observed a DNA accessibility enrichment and acetylation of Histone 3 at the promotor region of our candidates, correlating with DNase sensitivity hotspots in BM mesenchymal cells that reinforce the prediction of the expression windows. In particular, for Mlinc.28428.2, the transcript observed with long reads sequencing corresponded to the prediction made with short reads. It was also supported by Mlinc.28428.1, a variant that differ by the absence of the second exon. Similar characteristics were observed for Mlinc.128022, which also produced two variants with a different organisation of 5 exons. The two other candidates, Mlinc.89912.1 and Mlinc.64225.1 are mono-exonic. Mlinc.89912.1 occurs at the close proximity of FGF5 3' end, in reverse orientation at this locus. For this reason, the different epigenomic features could not be attributed with certainty to the Mlinc. For Mlinc.64225.1, the sequence is longer than the *ab initio* short read prediction. KRTAP1-5, HOTAIR and SMILR, selected for their good expression profile, were also covered by long reads (Data not shown).

2.18 High-throughput investigation of a marker's specificity by specific k-mers search

A marker can only be considered specific within the limits of the diversity of samples used for its study. Considering the growing number of cell/tissues and transcriptional profiles, it is essential to probe the limits of a chosen biomarker against these various cell types. Most of published analyses highlighting new potential biomarkers of MSCs or fibroblasts have been restricted to a comparison between only few cell types and, as discussed, commonly described markers are not strictly distinctive. In order to assess the expression of Mlincs candidates in a large number of samples, we extracted specific 31nt k-mers from each of their sequences, as previously described [49]. These simplified but candidate-specific (oligonucleotide-like) probes allow a simple and fast presence/absence search to be made on large-scale cohorts and a direct quantification in raw fastq data. The k-mers were quantified in ENCODE human RNAseq database, including "primary cells" and "*in vitro* differentiated cell" categories (Table S2). Particularly, as the bibliography suggests that MSCs can also express phenotypic characteristics of endothelial, neural, smooth

muscle cells, skeletal myoblasts and cardiac myocytes, RNAseq samples from this mesodermal origin were tested.

With ISCT positive markers, we observed an expected expression profile that recapitulate previous biological studies, particularly the high expression of Endoglin (ENG, CD105) in endothelial cells (Figure S4) and the overexpression of NT5E (CD73) in epithelial and endothelial cells (Figure S5). Interestingly, their expression varied among MSCs cell sources : NT5E (CD73) was strongly enriched in adipose and BM derived MSCs, and THY1 (CD90) in umbilical cord derived MSCs (UC-MSCs) (Figure S6). We next analysed the expression profile using our candidate annotated genes Mlinc specific k-mers (Figure 4). The specific k-mers search supported the stated expression profile of Mlincs previously shown : our Mlinc candidates were positive in MSCs and displayed weak or absent expression in cells of ectodermal lineage, hematopoietic or endothelial origins.

However the high throughput and naive quantification in the ENCODE cohort made it possible to extend the observation of this absence of expression into cell types not previously studied. Moreover, this diversity showed that the expression of most of the candidates, contrarily to positive markers of the ISCT, were exclusive of cells with mesodermal origin. All candidates were expressed in at least one type of fibroblasts and differentially present in other mesodermal cell types. For the 4 selected Mlincs, they shared: (i) a systematic and strong expression in cell types like skin fibroblasts, and cells derived from reservoir of mesenchymal progenitors (muscle satellite cells or dermis papilla cells), (ii) a regular over-expression in regular cardiac myocytes, (iii) a significant and irregular expression in smooth muscle cells. The ENCODE cohort containing MSCs of different origins, we can therefore observe that the Mlincs show differences of expression depending of the tissular origin, these candidates being mainly expressed in two MSCs types. The results permitted the classification of our Mlincs according to observed specificity, from the most promising to the least restricted profile : Mlinc.28428.2 is expressed in Ad and BM derived MSCs. It is the candidate with the clearest absence of expression in non-mesodermal cells, and with the poorest relative expression in Smooth Muscle Cells (SMC). Mlinc.128022.2 is expressed in Adipose and Bone Marrow MSCs, and particularly in preadipocytes and muscle cells : (myoblasts, myocytes and myotubes). Mlinc.89912.1 is principally expressed BM-MSCs and less in UC-MSCs and AD-MSCs, but shows expression in epithelial and endothelial cell. Finally, Mlinc.64225.1 differs from other Mlincs as it is also strongly expressed in keratinocytes, HSC, and epithelial cells (Figure S7). Its expression in critical non-MSC types has led us to retain the three other Mlincs for further investigations.

2.19 RT-qPCR mimics the *in silico* prediction and deciphers multiple transcript variants

To confirm the specificity of selected Mlincs' expression experimentally, we performed RT-qPCR on a set of 80 RNA preparations from different primary cell (Figure 4C). These includes MSCs from Bone Marrow (BM), Adipose (Ad) and Umbilical Cord (UC), fibroblasts of different tissue origins, IPS cells, neural stem cells, myoblasts, HUVECs and hepatocytes (Table S6). RT-qPCR and amplicon sequencing using sets of specific primers (Table S7) confirmed different predicted forms of the Mlincs candidates in BM-MSCs (Figure S8). We designed two primer pairs for both Mlinc.128022 variants to validate the existence of first splice, and two pairs for Mlinc.28428 variants, one overlapping the second exon and another corresponding to a splice between first and third exon. All variations captured by the primer designs were quantified, suggesting that all these different variations predicted *in silico* exist biologically in MSCs. We confirmed most of the

expression profiles obtained by k-mers quantification using RT-qPCR, notably the specificity of expression dependency on the MSCs tissular origin: over expression of Mlinc.28428 and 128022 in BM- and Ad-MSCs. Nevertheless, few exceptions, such as Mlinc.89912.1, present an enrichment in UC-MSCs not found in k-mers quantification. Moreover, the restricted expression to cells of mesodermal origin is replicated in our RT-qPCR results. We obtained similar observations with annotated candidates : overexpression of KRTAP1-5 and SMILR in BM-MSCs specifically, and of HOTAIR in UC- and BM-MSCs.

2.20 *In silico* prediction of lncRNAs interactions and functions

The relative specificity of selected Mlincs for mesenchymal cells could be an indication of their roles in MSCs function. The prediction of their possible function could therefore suggest their suitability as markers of MSCs' function potential. To this end, we explored assumptions on the function of Mlinc.28428.2, Mlinc.128022.2 and Mlinc.89912.1 candidates using different published methods. We first used bioinformatic tools based on machine learning and deep learning to decipher general characteristics of our candidates : FEELnc to assess coding potential, tarpMir to decipher "miRNA sponge" function, and LncADeep to analyse potential interactions with proteins. Only two of the 35 selected Mlincs and none of the 3 selected Mlincs with validated specificity were revealed as potentially coding RNAs with the majority being predicted as non-coding by FEELnc (33/35). None candidate had more than five target sites for a given miRNA, indicating a low probability of a "miRNA sponge" activity (Table S7). For the 3 retained Mlincs, predicted interacting proteins by LncADeep were submitted to Reactome (Table S8).

We noted a predicted interaction between Mlinc.28428.2 and Betacatenin (CTNNB1) as part of apoptosis-linked modules, 5'-3' Exoribonuclease 1, component of the CCR4-NOT complex, mRNA Decapping Enzyme 1B as part of the mRNA decapping and decay pathways. The interaction was also predicted with different mediators of RNA polymerase II transcription subunits (MED), ATP Binding Cassette Subfamily B Members as part of the PPARA activity linked to ER-stress [50], and Proteasome subunits for intracellular transport, response to hypoxia and cell cycle modules. Mlinc.128022 could interact with important genes like THY1 (CD90), NRF1 (mitochondria metabolism) with no module clearly highlighted. Mlinc.89912 could interact with tubulins, UBB (ubiquitin B), SMG6 nonsense mediated mRNA decay factor and ribosomes subunits (RPSX) proteins, (RPL24) for nonsense Mediated Decay (NMD), PINK1 (mitophagy) and finally MGMT, as part of the MGMT mediated DNA damage reversal module.

We further quantified expression of candidates by counting their specific k-mers in the entire FANTOM6 set of 154 Known-Downed (KD) annotated lncRNAs in human dermal fibroblasts (<https://doi.org/10.1101/700864>, not peer-reviewed). We selected the KD experiments where expression of the Mlincs was statistically differential when compared with controls. Particular attention was paid to KD lncRNAs with reported function(s) in bibliography, and to KD lncRNAs overlapping a gene with reported functions. Mlinc.28428.2 is down-regulated when JPX, SERTAD4-AS1, BOLA3-AS1, and SNRPD3 are KD, and over-expressed with the KD of PTCHD3P1, ERVK3.1, MEG3, among other lncRNAs without reported function (Figure 5A). Interestingly, interactions between p53 pathway and JPX [51], SNRPD3 [52] and MEG3 [53, 54] respectively, have been previously reported. All these features converge on the hypothesis of a link between the function of Mlinc.28428, stress response, senescence and cellular maintenance. The implications of BOLA3 [55, 56]) and PTCHD3P1 [57] in mitochondria homeostasis and glycolysis, the role of BOLA3 in stress response [58], the status of SERTAD4 as target of the YAP/TAZ pathway [59], vital pathway of stress response [60], and the role of MEG3 in aging [61], all reinforce

this hypothesis.

Mlinc.128022.2 is down-regulated with the KD of FOXN3-AS1, A1BG-AS1, CD27-AS1, and FLVCR1-AS1 (Figure 5B). FOXN3 seems to be more than a regulator of cell cycle, it is also described as a regulator of osteogenesis in different cases of defective craniofacial development [62, 63]. Moreover, the reported over-expression of FOXN3 during the early stages of MSCs osteodifferentiation [64], and downregulation of CD27-AS1 in MSCs of donors with bone fracture [65], allow us to hypothesise a possible function of Mlinc.128022 in bone remodelling and osteogenesis. In addition, both A1BG-AS1 and FLVCR1-AS have an influence in osteogenesis and cell differentiation. A recent study showed that A1BG-AS1 interacts with miR-216a and SMAD7 in suppressing hepatocellular carcinoma proliferation [66], both partners having an important role in the positive regulation of osteoblastic differentiation in mice [67] [68]. FLVCR1 participates to resistance of oxydative stress by heme exportation in mouse MSCs [69], iron metabolism being closely linked with bone homeostasis, formation [70] and cell differentiation [71].

Finally, Mlinc.89912.1 is down-regulated after the KD of NEAT1-1 and PCAT6, and over-expressed when MFI2.AS1, CDKN2B.AS1 (or ANRIL) and MKLN1.AS2 are KD (Figure 5C). The manifest relations between cell proliferation and CDKN2B-AS1 [72, 73], MFI2 [74], MFI.AS1 [75], PCAT6 [76] and NEAT1 [77, 78], an combination with the between ones and DNA damage repair response, [79, 80] reinforces the prediction of a role of Mlinc.89912 in these mechanisms. Moreover, we explored RNAseq from chromatin, nucleus and cytoplasm subcellular compartments of fibroblastic cells in the FANTOM6 Dataset. Mlinc.28428 and Mlinc.128022 are enriched in at least cytoplasm, whereas Mlinc.89912 is enriched in free nucleus fraction suggesting interaction with nuclear component (Figure 5C).

2.21 The single cell RNAseq: an emergent level of completion in marker search

We analysed the single cell RNAseq data from 26 071 adipose MSCs (Ad-MSCs) to assess the heterogeneity of the 3 Mlincs and to explore their expression at the single-cell level. No clear correlation between cell cycle and expression of our Mlincs was identified (Figure S9). We observed a high variability of the number of cells expressing the markers (Threshold ≥ 0.1). 11 927/26 071 were Mlinc.28428-positives, 4944 were Mlinc.128022-positives, and 404 were Mlinc.89912-positives. For each Mlinc, we performed a differential test to decipher genes differentially expressed in Ad-MSCs Mlinc-positive and Mlinc-negative cells.

We found that Mlinc.28428-positive cells under-expressed H19 and PI16 (Figure 5A). These genes, that present a diversity of functions, are involved in stress mechanisms (oxydative response and shear stress), inflammation in fibroblasts and MSCs, and senescence pathways [81, 82, 83, 84]. Despite the low number of differentially expressed genes in Mlinc.28428-positive cells, their functional behaviour and their known targets suggest a pathway linked to stress response and senescence establishment that reinforce our previous assumptions on Mlinc.28428 function.

Mlinc.128022-positive cells are enriched in FTH1, TPM2, FTL and CD24 and present a lower expression in HMGN2, HMGB1, ODC1, PTTG1, BIRC5, EIF5A, MKI67, UBE2S, FGF5, HAS2-AS1 (Figure 5B). A significant portion of these genes are linked to osteogenic properties of MSCs as previously observed with FANTOM analysis. The Mlinc.128022-positive cells have an increased expression of ferritin (light and heavy chains), major actor in iron metabolism in osteoblastic cell line [85], that is also involved in osteogenic differentiation [86] and osteogenic calcification [87]. Two genes, enriched in Mlinc.128022-positive cells are positively linked to the osteogenic

differentiation potential of MSCs : the tropomyosin 2 (TPM2), downregulated when hMSCs were cultured in OS medium for the induction of osteoblasts at the calcification phase [88], and CD24 a membrane antigen recently proposed as a new marker for the sub-fraction of notochordal cells with increased differentiation capabilities [89]. In addition ODC1, under-represented in Mlinc.128022-positive cells, inhibited the MSCs osteogenic differentiation [90, 91]. Finally, the decrease of FGF5, MKI67, BIRC5 (Survivin) and PTTG1 (securin) expressions, all linked to proliferation active phases of cell cycle, tend to show cell with arrested cell cycle. These data suggest that the expression profile of Mlinc.128022 positive cells indicate a subpopulation of undifferentiated osteogenic progenitors, probably in senescence or quiescence.

Mlinc.89912-positive cells are enriched in FGF5 and HIST1H4C (Figure 5C). FGF5 is a protein with mitogenic properties, identified as an oncogene, that facilitates cell proliferation in both autocrine [92] and paracrine manner [93]. HIST1H4C, the Histone Core number 4, is a cell cycle-related gene. Modification of histone H4 (post-transcriptional or mutation) has been highlighted as important for Non-Homologous End-Joining (NHEJ) in yeast [94]. Its mutation cause genomic instability, resulting in increased apoptosis and cell cycle progression anomalies in zebrafish development. It reinforces our assumptions that Mlinc.89912 has a role in cell proliferation and DNA damage repair. In conclusion, the single cell RNAseq analysis enabled the observation of different features that characterise the phenotype of Mlincs positive cells and reinforced hypotheses on their functions previously observed through k-mers quantification.

2.22 K-mers analysis of markers in functional cell situation

Previously, we have presented a number of strategies to formulate hypotheses on the functions of an unannotated lncRNAs, suggesting directions of future experimental investigations. To evaluate the relevance of these strategies, we sought to quantify with specific k-mers search the expression of our Mlincs in MSCs in different conditions, linked to above mentioned findings: stress and senescence for Mlinc.28428.2, osteodifferentiation for Mlinc.128022.2 and cell cycle/proliferation for Mlinc.89912. We downloaded RNAseq data corresponding to the above-mentioned focus, described in Table S4.

As shown in Figure 6, we observed a statistically relevant increase of Mlinc.28428 expression in MSCs under replicative stress and in MSCs with CrisprCas9 depletion of genes with important role against senescence. In the Wang et al. study [95], MSCs senescence was observed with the KO of ATF6 and the stress induced with tunicamycin (endoplasmic reticulum stress) and late passage (replicative stress). Mlinc.28428 expression increased with tunicamycin treatment, late passage and ATF6 KO. The highest increase is observed in ATF6 KO MSCs associated with late passage condition.

In Fu et al. study [96], YAP, but not TAZ, was found to safeguard MSCs from cellular senescence as shown by KO experiments. Interestingly, YAP KO, but not TAZ, significantly increases the expression of Mlinc.28428.2. This would lead us to conclude that Mlinc.28428 is overexpressed in senescence and stress conditions, suggesting a role in one or both of these phenomena.

The change in Mlinc.128022 expression is strictly linked to osteodifferentiation conditions. Mlinc.128022 expression shows a relevant increase in MSCs exposed to fungal metabolite Cycloheximide (CytoD). The cytoD is reported as a osteogenic stimulant in the concerned study [97]. Moreover, no expression variation was observed between MSCs and MSC-derived adipocytes from Wang et al. study, implying a role in adipodifferentiation. Agrawal Singh et al. have studied osteogenic MSCs differentiation [29], with a similar increase of Mlinc.128022 being observed after ten days.

We then quantified the expression of Mlinc.89912 in a study that compare proliferating MSCs versus confluent MSCs [98, 99]. Our candidate was clearly over-expressed in proliferating cells, validating its capacity to mark the MSCs in proliferation. Moreover, its expression was not statistically modified when MSCs were exposed to EGF with pro-mitotic capabilities [100]. However Mlinc.89912 expression was reduced when IWR-1, an inhibitor of Beta-catenin nuclear translocation, that reduced the proliferation of MSCs, was added to the medium. The functional domains of these genes is summarised in table 1 and confirm the potential functional role suggested from FANTOM data: stress-related pathways for Mlinc.28428, MSCs differentiation with a presumed orientation in osteo-progenitors for Mlinc.128022 and a more restricted role in proliferation and DNA repair for Mlinc.89912.

3 Discussion

With recent evolution of omics analysis, the landscape of biomarkers has been extended beyond known genes to the exploration of the unexplored transcriptome. This potential has been assessed in pathological conditions but to a lesser extent in cell-specific conditions, where this new pool of potential markers could be used to identify less well-characterised cells and hence predict their function. In this article, we propose an integrated procedure and strategies to identify the best markers (annotated or not) in a cell-specific condition, and predict their potential functions, primarily from RNA sequencing data (Figure 1). RNAseq facilitates the creation of large lncRNA catalogues [8, 101] through the total catalogue of lncRNAs remains incomplete given the diversity of biological entities and lncRNAs specific expression in non-pathological, cell-specific conditions. The creation of a "home-made" catalogue associated with a specific condition remains the best way to assess the full diversity of potential biomarkers in a cell, rather than resorting to a global catalogue made from diverse samples. To give an idea of the completeness of such a focused lncRNA catalogue when compared to a global one, Jiang et al. recently published "an expanded landscape of human long non-coding RNA" with 25 000 new lncRNAs from normal and tumor tissues, whereas in our focused analysis only 50% of our 35 selected MSClinc can be found in this collection [101].

Futhermore, providing new candidates of good quality to improve lncRNA collection remains a complex task. As could be expected, the raw catalogue in our study contains predictions of disparate quality observed with a large number of mono-exonic transcripts. Without any filter, *ab initio* methods are insufficient to adequately reconstruct full length transcripts. The usage of long-read sequencing has been particularly effective in helping to validate our predictions. Given the benefits of full-length RNA sequence, long-read RNAseq should become the standard for lncRNA validation. A specific lncRNA can be the one presenting the most relevant properties after *in silico* analysis. The first task remain the identification of the more specific markers for a given cell type, task that present differences from classic comparative analysis. The MSCs markers proposed in the past were determined through a simple comparison between MSCs of a certain origin with negative cell whose types are either unique or few in number.

Historically, MSCs have been compared to bone marrow haematopoietic stem cells. However, our initial RNAseq analysis revealed that all potential MSCs markers proposed in the past are expressed in at least one other non-mesenchymatous cells type, and so do not constitute exclusive MSCs markers at the transcriptome level. Even if all cell types cannot be investigated, the diversity of the negative cell set is a critical criterion in selecting the most specific transcripts. In keeping with this idea, we then restricted the list of potential biomarkers with an enrichment step based

Table 1: Summary of functional investigation results

Mlinc	Predicted RNA-Protein interactions (lncADeep)	subcompartment enrichment	FANTOM6 expr. changes	Diff. genes in positive cells	K-mers investigations
Mlinc.28428	apoptosis, mRNA decay, PPARA activity, intracellular transport, response to hypoxia and cell cycle	Chromatin, cytoplasm	BOLA3-AS1, JPX, SERTAD4-AS1, PTCHD3P1, ERVK3.1, SNRPD3, MEG3	H19 - PI16 -	Stress, senescence
Mlinc.128022	THY1, NRF1	Chromatin, cytoplasm	FOXN3-AS1, A1BG-AS1, CD27-AS1, FLVCR1-AS1	FTH1, TPM2, FTL, CD24, HMGN2, HMGB1, ODC1, PTTG1, BIRC5, EIF5A, MKI67, UBE2S, FGF5, HAS2-AS1	Osteodiff., stress
Mlinc.89912	MGMT-mediated DNA damage reversal, Nonsense Mediated Decay, Tubulin metabolism	nucleus (free), cytoplasm	NEAT1.1, PCAT6, MFI2.AS1, MKLN1.AS2, CDKN2B.AS1	FGF5 and HIST1H4C	Proliferation

on a differential expression comparing BM-MSCs to other cells including mainly stem cells, as well as differentiated cells of various lineages (lymphocytes, macrophages, primary chondrocytes, hepatocytes and neurons). In the enriched list, the overexpressed annotated genes contained members of MSC-related pathways as well as the ISCT markers. This result supported the MSCs characterisation made by the original authors [13], thus validating the identity of MSCs used for this RNAseq analysis with the currently defined criteria. The problem with classical differential

analysis used on diverse "non-MSc" group is that all the "non-MSc" group is considered to be homogeneous. As a result, candidates with positive expression in small cell groups could pass statistical test, creating false positives. For this kind of differential analysis, we propose to select the most discriminating transcripts by feature selection, a machine learning methodology, that reduces the number of non-discriminating candidates after selection. We used feature selection through Boruta, a method based on random forest, to retain the top 35 of the most relevant MSCs signature for annotated genes, Mlincs and Mloancs separately. Putting aside our initial focus on unannotated lncRNAs, different annotated lncRNAs or coding genes with interesting profiles were also selected by feature selection : among them, KRTAP1-5 have been exclusively studied in BM-MSCs [102], where its preferential expression was validated by our results. These discoveries can bring new features concerning these genes and suggest directions for future investigations concerning their impact on the MSCs.

However, a marker is classically considered as specific on condition that its positive expression cannot be observed in any other cell type. Therefore, the expression of these potential markers should be explored in an entire RNAseq database to further validate its specificity. The exploration of a wide set of RNAseq data as proposed by ENCODE including a diversified set of primary and stem cells could support or invalidate the specificity of potential markers. In order to assess the expression of Mlincs candidates in a large number of samples, we used a signature for each candidate, extracting specific 31nt k-mers from their sequences. The specific k-mers extraction was made using Kmerator software. These k-mers were then quantified in the ENCODE human RNAseq database. The new and simplified procedure based on k-mers counting and large scale RNAseq exploration has the following advantages: i) a direct textual search that requires less time and CPU resources than classical methods, ii) a restricted set of lncRNAs supported by different results in the biological (wet) and *in silico* level (RNAseq data). The counterpart of the extensive vision of marker expression is that we observe a limit of specificity among our best candidates. We observed expression in fibroblasts, in close primary cells of common embryonic origin like smooth muscle cells (SMC) and other tissue-specific fibroblastic cells. Other tissue resident fibroblastic cells like skeletal muscle satellite cells, pre-adipocytes, and fibroblasts from different sources, especially dermis, express our selected MSClincs markers. The question of the differences between MSCs and related cell types is crucial to the issue. Specifically, the differences between MSCs and fibroblasts remain a subject of debate [103, 12]. According to the ISCT statement, no phenotypical differences have been reported between fibroblasts of different sources and adult MSCs [104], suggesting a hypothesis of a uniform cell type that show functional variation depending on the tissue source. Our results support this idea: distinguishing MSCs from fibroblast with only few positive markers remains a complicated task.

Moreover we observe low to medium expression of our candidates in close cell type from the same embryonic origin such as muscular cells and smooth muscle cells (SMC). This could be due to a shared phenotype between cells with close embryonic origin. Common markers between MSCs and SMCs have already been described. Notably, MSCs can express similar levels of SMC markers such as alpha-actin [105, 106]. Moreover Kumar et al. [107] determined that MSCs, pericytes, and SMCs could have the same mesenchymo-angioblast progenitor and that SMCs share a certain plasticity with MSCs as they can be differentiated in chondrocyte-like and beige adipocytes or myofibroblasts. However, a lot of cell types in ENCODE have not been actively sorted by expression of their respective surface markers, and fibroblast contamination is a classical feature in primary cell culture. We should not therefore exclude the possibility of fibroblast contamination when investigating marker for MSCs by bulk omics technology. Given this, single-cell RNAseq could be the best solution to identify the source of marker expression in counterpart cells.

To conclude, our extensive cell type comparison shows that the discovery of a marker of MSCs as distinct cell type is not plausible. After deepening our own research on MSCs biomarkers at the annotated and unannotated level, we were unable to find a marker that could simultaneously i) distinguish MSCs to close or homologous cell types (fibroblasts, satellite cells, SMCs) ii) be present in all MSCs types iii) distinguish MSCs from more characterised cell types (Hematopoietic lineage, neurones etc). Our results suggest, like other studies, a strong proximity between MSCs, fibroblast and mesodermal cell types.

More than a marker of MSCs, candidates extracted by our method could be used to explore important features in MSCs biology and therefore warrant investigation into their function, assuming that the specificity of RNA for a cell type can highlight its importance in cell activity. Even if the functional invalidation stands as the principal method to efficiently determine the function of a lncRNA, its expression and co-expression with known genes can potentially characterise a function or an intrinsic state of a cell type, particularly for MSCs with reported diversity of states and function (ex : differentiation, immunomodulation, senescence, proliferation...). In our opinion, it is vital that during the creation of a catalogue of lncRNAs, a restricted set of selected biomarkers should be studied more intensively, both in term of specificity and functions. Assumptions on functional domains, where lncRNAs could act, could increase the relevance and visibility of discovered lncRNAs, and far from the bioinformatics implications, encourage future biological investigations. We decided to investigate the three selected MSClincs, validated by k-mers search, RT-qPCR and long-read sequencing, in term of biological impact with complementary *in silico* experimental approaches. We propose to use different *in silico* strategies, depending on the amount and diversity of the available data. The analysis confirms the non-coding potential of candidates and indicates a low probability of "miRNA sponge" activity. However, protein potential interaction results give interesting paths that were then investigated by complementary exploration. The k-mers quantification permits a naive high throughput exploration of numerous RNAseq data, simultaneously exploring potential functions and specificity to assess their potential. Instead of different cells, each candidate's expression was quantified in MSCs in different experimental conditions. FANTOM6 data recently offered a pilot about lncRNAs functional investigation, with a high-throughput invalidation of 154 lncRNAs and coding genes in fibroblasts and their RNAseq counterpart added to phenotypical observations. The utilisation of co-expressions between knock-out genes and candidates lncRNAs remains an efficient way to decipher lncRNAs function, provided number of KD genes is high. Moreover, the availability of recent single cell data of MSCs have been a good complement to lncRNAs functional investigation.

Using scRNAseq from Ad-MSCs [19], we observed that our markers are not expressed in all cells but constitute different subpopulations with different levels of rarity in Ad-MSCs. FANTOM6 and single-cell analysis could permit tracing three components of these states : stress inducible cells, lineage committed osteogenic progenitors and proliferating cells. Globally, we observed a global concordance of the results between the different strategies used for functional prediction. Mlinc.28428 has concomitant expression with genes related to the stress response pathway. Mlinc.28428 could be a good target for treatment to study the senescence process, age pathologies or stress response. Mlinc.128022 potentially interacts with THY1 (CD90) and has co-occurrences with genes linked to osteoprogenitors and cell differentiation. The k-mers search highlights its participation in MSCs' osteo-differentiation. Finally, Mlinc.89912 potentially interacts with damage repair and RNA decay, and tubulin metabolism, all linked to cell proliferation and cell cycle. Moreover, the subcompartment enrichment corresponds to this prediction: Mlinc.89912.1 is the only candidate to have possible interactions with DNA-repair system, a hypothesis corresponding to his observed enrichment in the nucleus. A final selection of bulk RNAseq of MSCs in specific biological con-

ditions allowed confirmation of our initial assumptions, showing that the different strategies we propose could be used to give relevant indication of the lncRNAs' functions. These results show that a lncRNA selected by its expression specificity has a high probability of being part of a functional mechanism.

In conclusion, we have predicted genes and lncRNAs enriched in MSCs and proposed several selection steps including feature selection (machine learning), large scale signature search, RT-qPCR validation, *in silico* tools and single cell analysis. We present the application of a new way of quantification in RNAseq : The specific k-mers search could be used as a naive information in lncRNA catalogue creation. The strategies presented here are transferable to other cell types and different studies while the specificity and functional assumption present a significant potential in long non-coding transcriptome exploration. We present three lncRNAs markers of bone marrow and adipose MSCs that passed all selection steps and present interesting features: Mlinc.28428.2, Mlinc.128022.2 and Mlinc.89912.1. These markers could be used by the scientific community as potential targets for functional biological experiments on MSCs, with pre-indications of potential functions to orientate the experiments, and finally initiate the objective of transition between informatical problematics and cell biology.

4 Funding

Grant information: this work was supported by the Agence Nationale de la recherche for the projects "Computational Biology Institute" and "Transipedia" '[grant numbers 18-CE45-0020-02, ANR-10-INBS-09]' and the Canceropole Grand-Sud-Ouest Trans-kmer" project '[grant number 2017-EM24]'.

Acknowledgements

We thank for their generous gifts, G.Carnac for myoblasts, M.Le Quintrec-Donnette for HUVECs, E. Sanchez for dermal fibroblasts, D. Noel and ML. Vignais for mesenchymal stromal cells, C. Crozet for IPS, S. Gerbal and M. Daujat for hepatocytes. We thank Philippe Clair for his advice on qPCR, the qPHD platform, Montpellier GenomiX and Jean-Marc Holder (SeqOne) for text corrections.

4.0.1 Conflict of interest statement.

The authors declare that they have no competing interests.

References

- [1] Gloss, B. S. and Dinger, M. E. (January, 2016) The specificity of long noncoding RNA expression. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1859**(1), 16–22.
- [2] Meseure, D., Drak Alsibai, K., Nicolas, A., Bieche, I., and Morillon, A. (2015) Long Noncoding RNAs as New Architects in Cancer Epigenetics, Prognostic Biomarkers, and Potential Therapeutic Targets. *BioMed Research International*, **2015**.

- [3] Bouckenheimer, J., Assou, S., Riquier, S., Hou, C., Philippe, N., Sansac, C., Lavabre-Bertrand, T., Commes, T., Lematre, J.-M., Boureux, A., and Vos, J. D. (September, 2016) Long non-coding RNAs in human early embryonic development and their potential in ART. *Human Reproduction Update*,.
- [4] Li, L. and Chang, H. Y. (October, 2014) Physiological roles of long noncoding RNAs: Insights from knockout mice. *Trends in cell biology*, **24**(10), 594–602.
- [5] Dhamija, S. and Diederichs, S. (2016) From junk to master regulators of invasion: lncRNA functions in migration, EMT and metastasis. *International Journal of Cancer*, **139**(2), 269–280.
- [6] Li, X. and Li, N. (December, 2018) LncRNAs on guard. *International Immunopharmacology*, **65**, 60–63.
- [7] Morillon, A. and Gautheret, D. (June, 2019) Bridging the gap between reference and real transcriptomes. *Genome Biology*, **20**.
- [8] Uszczyńska-Ratajczak, B., Lagarde, J., Frankish, A., Guig, R., and Johnson, R. (September, 2018) Towards a complete map of the human long non-coding RNA transcriptome. *Nature reviews. Genetics*, **19**(9), 535–548.
- [9] James, A. R., Schroeder, M. P., Neumann, M., Bastian, L., Eckert, C., Gkbuget, N., Tanchez, J. O., Schlee, C., Isaakidis, K., Schwartz, S., Burmeister, T., von Stackelberg, A., Rieger, M. A., Gllner, S., Horstman, M., Schrappe, M., Kirschner-Schwabe, R., Brggemann, M., Mller-Tidow, C., Serve, H., Akalin, A., and Baldus, C. D. (January, 2019) Long non-coding RNAs defining major subtypes of B cell precursor acute lymphoblastic leukemia. *Journal of Hematology & Oncology*, **12**.
- [10] Liu, X., Ma, Y., Yin, K., Li, W., Chen, W., Zhang, Y., Zhu, C., Li, T., Han, B., Liu, X., Wang, S., and Zhou, Z. (June, 2019) Long non-coding and coding RNA profiling using strand-specific RNA-seq in human hypertrophic cardiomyopathy. *Scientific Data*, **6**(1), 1–7.
- [11] Lv, F.-J., Tuan, R. S., Cheung, K. M., and Leung, V. Y. (June, 2014) Concise Review: The Surface Markers and Identity of Human Mesenchymal Stem Cells. *STEM CELLS*, **32**(6), 1408–1419.
- [12] Soundararajan, M. and Kannan, S. (December, 2018) Fibroblasts and mesenchymal stem cells: Two sides of the same coin?. *Journal of Cellular Physiology*, **233**(12), 9099–9109.
- [13] Dominici, M., Le Blanc, K., Mueller, I., Slaper-Cortenbach, I., Marini, F., Krause, D., Deans, R., Keating, A., Prockop, D., and Horwitz, E. (2006) Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy*, **8**(4), 315–317.
- [14] Fitzsimmons, R. E. B., Mazurek, M. S., Soos, A., and Simmons, C. A. (August, 2018) Mesenchymal Stromal/Stem Cells in Regenerative Medicine and Tissue Engineering. *Stem Cells International*, **2018**.
- [15] Olsen, T. R., Ng, K. S., Lock, L. T., Ahsan, T., and Rowley, J. A. (June, 2018) Peak MSCAre We There Yet?. *Frontiers in Medicine*, **5**.

- [16] Tye, C. E., Gordon, J. A. R., Martin-Buley, L. A., Stein, J. L., Lian, J. B., and Stein, G. S. (March, 2015) Could lncRNAs be the missing links in control of mesenchymal stem cell differentiation?. *Journal of Cellular Physiology*, **230**(3), 526–534.
- [17] Kalwa, M., Hnzelmann, S., Otto, S., Kuo, C.-C., Franzen, J., Jousen, S., Fernandez-Rebollo, E., Rath, B., Koch, C., Hofmann, A., Lee, S.-H., Teschendorff, A. E., Denecke, B., Lin, Q., Widschwendter, M., Weinhold, E., Costa, I. G., and Wagner, W. (December, 2016) The lncRNA HOTAIR impacts on mesenchymal stem cells via triple helix formation. *Nucleic Acids Research*, **44**(22), 10631–10643.
- [18] Song, W., Gu, W., Qian, Y., Ma, X., Mao, Y., and Liu, W. (2015) Identification of long non-coding RNA involved in osteogenic differentiation from mesenchymal stem cells using RNA-Seq data. *Genetics and Molecular Research*, **14**(4), 18268–18279.
- [19] Liu, X., Xiang, Q., Xu, F., Huang, J., Yu, N., Zhang, Q., Long, X., and Zhou, Z. (February, 2019) Single-cell RNA-seq of cultured human adipose-derived mesenchymal stem cells. *Scientific Data*, **6**, 190031.
- [20] Peffers, M. J., Collins, J., Fang, Y., Goljanek-Whysall, K., Rushton, M., Loughlin, J., Proctor, C., and Clegg, P. D. (2016) Age-related changes in mesenchymal stem cells identified using a multi-omics approach. *European Cells & Materials*, **31**, 136–159.
- [21] Philippe, N., Salson, M., Commes, T., and Rivals, E. (2013) CRAC: an integrated approach to the analysis of RNA-seq reads. *Genome Biology*, **14**, R30.
- [22] Perte, M., Perte, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (March, 2015) StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, **33**(3), 290–295.
- [23] Quinlan, A. R. and Hall, I. M. (March, 2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**(6), 841–842.
- [24] Ilt, N. E. and Ponting, C. P. (September, 2013) Predicting long non-coding RNAs using RNA sequencing. *Methods (San Diego, Calif.)*, **63**(1), 50–59.
- [25] Li, H. (July, 2016) Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics*, **32**(14), 2103–2110.
- [26] Bray, N. L., Pimentel, H., Melsted, P., and Pachter, L. (May, 2016) Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, **34**(5), 525–527.
- [27] Pimentel, H., Bray, N. L., Puente, S., Melsted, P., and Pachter, L. (July, 2017) Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods*, **14**(7), 687–690.
- [28] Kursa, M. B., Jankowski, A., and Rudnicki, W. R. (December, 2010) Boruta - A System for Feature Selection. *Fundam. Inf.*, **101**(4), 271–285.
- [29] Agrawal Singh, S., Lerdrup, M., Gomes, A.-L. R., van de Werken, H. J., Vilstrup Johansen, J., Andersson, R., Sandelin, A., Helin, K., and Hansen, K. PLZF targets developmental enhancers for activation during osteogenic differentiation of human mesenchymal stem cells. *eLife*, **8**.

- [30] Yang, C., Yang, L., Zhou, M., Xie, H., Zhang, C., Wang, M. D., and Zhu, H. (November, 2018) LncADeep: an ab initio lncRNA identification and functional annotation tool based on deep learning. *Bioinformatics*, **34**(22), 3825–3834.
- [31] Ding, J., Li, X., and Hu, H. (September, 2016) TarPmiR: a new approach for microRNA target site prediction. *Bioinformatics*, **32**(18), 2768–2775.
- [32] Wucher, V., Legeai, F., Hdan, B., Rizk, G., Lagoutte, L., Leeb, T., Jagannathan, V., Cadieu, E., David, A., Lohi, H., Cirera, S., Fredholm, M., Botherel, N., Leegwater, P. A. J., Le Bguec, C., Fieten, H., Johnson, J., Alfdi, J., Andr, C., Lindblad-Toh, K., Hitte, C., and Derrien, T. (May, 2017) FEELnc: a tool for long non-coding RNA annotation and its application to the dog transcriptome. *Nucleic Acids Research*, **45**(8), e57–e57.
- [33] Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (June, 2019) Comprehensive Integration of Single-Cell Data. *Cell*, **177**(7), 1888–1902.e21.
- [34] Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (May, 2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*, **36**(5), 411–420.
- [35] Johnson, W. E., Li, C., and Rabinovic, A. (January, 2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, **8**(1), 118–127.
- [36] Hahne, F. and Ivanek, R. (2016) Visualizing Genomic Data Using Gviz and Bioconductor. In Math, E. and Davis, S., (eds.), *Statistical Genomics: Methods and Protocols*, Methods in Molecular Biology pp. 335–351 Springer New York New York, NY.
- [37] Djouad, F., Bony, C., Hupl, T., Uz, G., Lahlou, N., Louis-Plence, P., Apparailly, F., Canovas, F., Rme, T., Sany, J., Jorgensen, C., and Nol, D. (2005) Transcriptional profiles discriminate bone marrow-derived and synovium-derived mesenchymal stem cells. *Arthritis Research & Therapy*, **7**(6), R1304–R1315.
- [38] Kitzmann, M., Bonnieu, A., Duret, C., Vernus, B., Barro, M., LaoudjChenivesse, D., Verdi, J. M., and Carnac, G. (2006) Inhibition of Notch signaling induces myotube hypertrophy by recruiting a subpopulation of reserve cells. *Journal of Cellular Physiology*, **208**(3), 538–548.
- [39] Pichard, L., Raulet, E., Fabre, G., Ferrini, J. B., Ourlin, J.-C., and Maurel, P. (2006) Human Hepatocyte Culture. In Phillips, I. R. and Shephard, E. A., (eds.), *Cytochrome P450 Protocols*, Methods in Molecular Biology pp. 283–293 Humana Press Totowa, NJ.
- [40] Livak, K. J. and Schmittgen, T. D. (December, 2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2CT Method. *Methods*, **25**(4), 402–408.
- [41] Niazi, F. and Valadkhan, S. (April, 2012) Computational analysis of functional long noncoding RNAs reveals lack of peptide-coding capacity and parallels with 3 UTRs. *RNA*, **18**(4), 825–843.
- [42] Wang, Y., Xu, T., He, W., Shen, X., Zhao, Q., Bai, J., and You, M. (January, 2018) Genome-wide identification and characterization of putative lncRNAs in the diamondback moth, *Plutella xylostella* (L.). *Genomics*, **110**(1), 35–42.

- [43] Cagirici, H. B., Alptekin, B., and Budak, H. (September, 2017) RNA Sequencing and Co-expressed Long Non-coding RNA in Modern and Wild Wheats. *Scientific Reports*, **7**.
- [44] Salari, R., Aksay, C., Karakoc, E., Unrau, P. J., Hajirasouliha, I., and Sahinalp, S. C. (May, 2009) smyRNA: A Novel Ab Initio ncRNA Gene Finder. *PLoS ONE*, **4**(5).
- [45] Gu, Q., Tian, H., Zhang, K., Chen, D., Chen, D., Wang, X., and Zhao, J. (2018) Wnt5a/FZD4 Mediates the Mechanical Stretch-Induced Osteogenic Differentiation of Bone Mesenchymal Stem Cells. *Cellular Physiology and Biochemistry*, **48**(1), 215–226.
- [46] Diederichs, S., Tonniere, V., Mrz, M., Dreher, S. I., Geisbsch, A., and Richter, W. (October, 2019) Regulation of WNT5A and WNT11 during MSC in vitro chondrogenesis: WNT inhibition lowers BMP and hedgehog activity, and reduces hypertrophy. *Cellular and molecular life sciences: CMLS*, **76**(19), 3875–3889.
- [47] Bermeo, S., Vidal, C., Zhou, H., and Duque, G. (2015) Lamin A/C Acts as an Essential Factor in Mesenchymal Stem Cell Differentiation Through the Regulation of the Dynamics of the Wnt/-Catenin Pathway. *Journal of Cellular Biochemistry*, **116**(10), 2344–2353.
- [48] Chung, K.-M., Hsu, S.-C., Chu, Y.-R., Lin, M.-Y., Jiaang, W.-T., Chen, R.-H., and Chen, X. (February, 2014) Fibroblast Activation Protein (FAP) Is Essential for the Migration of Bone Marrow Mesenchymal Stem Cells through RhoA Activation. *PLoS ONE*, **9**(2).
- [49] Ruffl, F., Audoux, J., Boureux, A., Beaumeunier, S., Gaillard, J.-B., Bou Samra, E., Megarbane, A., Cassinat, B., Chomienne, C., Alves, R., Riquier, S., Gilbert, N., Lemaitre, J.-M., Bacq-Daian, D., Boug, A. L., Philippe, N., and Commes, T. (December, 2017) New chimeric RNAs in acute myeloid leukemia. *F1000Research*, **6**, 1302.
- [50] Krieken, S. E. v. d., Popeijus, H. E., Mensink, R. P., and Plat, J. (2017) Link Between ER-Stress, PPAR-Alpha Activation, and BET Inhibition in Relation to Apolipoprotein A-I Transcription in HepG2 Cells. *Journal of Cellular Biochemistry*, **118**(8), 2161–2167.
- [51] Delbridge, A. R. D., Kueh, A. J., Ke, F., Zamudio, N. M., El-Saafin, F., Jansz, N., Wang, G.-Y., Iminoff, M., Beck, T., Haupt, S., Hu, Y., May, R. E., Whitehead, L., Tai, L., Chiang, W., Herold, M. J., Haupt, Y., Smyth, G. K., Thomas, T., Blewitt, M. E., Strasser, A., and Voss, A. K. (April, 2019) Loss of p53 Causes Stochastic Aberrant X-Chromosome Inactivation and Female-Specific Neural Tube Defects. *Cell Reports*, **27**(2), 442–454.e5.
- [52] Siebringvan Olst, E., Blijlevens, M., de Menezes, R. X., van der MeulenMuileman, I. H., Smit, E. F., and van Beusechem, V. W. (May, 2017) A genomewide siRNA screen for regulators of tumor suppressor p53 activity in human nonsmall cell lung cancer cells identifies components of the RNA splicing machinery as targets for anticancer treatment. *Molecular Oncology*, **11**(5), 534–551.
- [53] Zhou, Y., Zhong, Y., Wang, Y., Zhang, X., Batista, D. L., Gejman, R., Ansell, P. J., Zhao, J., Weng, C., and Klibanski, A. (August, 2007) Activation of p53 by MEG3 Non-coding RNA. *Journal of Biological Chemistry*, **282**(34), 24731–24742.
- [54] Uroda, T., Anastasakou, E., Rossi, A., Teulon, J.-M., Pellequer, J.-L., Annibale, P., Pessey, O., Inga, A., Chillin, I., and Marcia, M. (September, 2019) Conserved Pseudoknots in lncRNA MEG3 Are Essential for Stimulation of the p53 Pathway. *Molecular Cell*, **75**(5), 982–995.e9.

- [55] Haack, T. B., Rolinski, B., Haberberger, B., Zimmermann, F., Schum, J., Strecker, V., Graf, E., Athing, U., Hoppen, T., Wittig, I., Sperl, W., Freisinger, P., Mayr, J. A., Strom, T. M., Meitinger, T., and Prokisch, H. (2013) Homozygous missense mutation in BOLA3 causes multiple mitochondrial dysfunctions syndrome in two siblings. *Journal of Inherited Metabolic Disease*, **36**(1), 55–62.
- [56] Yu QiuJun, Tai Yi-Yin, Tang Ying, Zhao Jingsi, Negi Vinny, Culley Miranda K., Pilli Jyotsna, Sun Wei, Brugger Karin, Mayr Johannes, Saggarr Rajeev, Saggarr Rajan, Wallace W. Dean, Ross David J., Waxman Aaron B., Wendell Stacy G., Mullett Steven J., Sembrat John, Rojas Mauricio, Khan Omar F., Dahlman James E., Sugahara Masataka, Kagiya Nobuyuki, Satoh Taiju, Zhang Manling, Feng Ning, Gorcsan John, Vargas Sara O., Haley Kathleen J., Kumar Rahul, Graham Brian B., Langer Robert, Anderson Daniel G., Wang Bing, Shiva Sruti, Bertero Thomas, and Chan Stephen Y. (May, 2019) BOLA (Bola Family Member 3) Deficiency Controls Endothelial Metabolism and Glycine Homeostasis in Pulmonary Hypertension. *Circulation*, **139**(19), 2238–2255.
- [57] Wang, J. and Li, K. (April, 2018) AB042. P013. lncRNAPTCHD3P1 enhances chemosensitivity of gemcitabine in pancreatic cancer and inhibits cancer cell proliferation and metastasis via inhibiting Warburg effect. *Annals of Pancreatic Cancer*, **1**(4).
- [58] Qin, L., Wang, M., Zuo, J., Feng, X., Liang, X., Wu, Z., and Ye, H. (April, 2015) Cytosolic Bola Plays a Repressive Role in the Tolerance against Excess Iron and MV-Induced Oxidative Stress in Plants. *PLoS ONE*, **10**(4).
- [59] Kitajima, S., Asahina, H., Chen, T., Guo, S., Quiceno, L. G., Cavanaugh, J. D., Merlino, A. A., Tange, S., Terai, H., Kim, J. W., Wang, X., Zhou, S., Xu, M., Wang, S., Zhu, Z., Thai, T. C., Takahashi, C., Wang, Y., Neve, R., Stinson, S., Tamayo, P., Watanabe, H., Kirschmeier, P. T., Wong, K.-K., and Barbie, D. A. (September, 2018) Overcoming Resistance to Dual Innate Immune and MEK Inhibition Downstream of KRAS. *Cancer cell*, **34**(3), 439–452.e6.
- [60] Raj, N. and Bam, R. (August, 2019) Reciprocal Crosstalk Between YAP1/Hippo Pathway and the p53 Family Proteins: Mechanisms and Outcomes in Cancer. *Frontiers in Cell and Developmental Biology*, **7**.
- [61] He, J., Tu, C., and Liu, Y. (2018) Role of lncRNAs in aging and age-related diseases. *AGING MEDICINE*, **1**(2), 158–175.
- [62] Schuff, M., Rssner, A., Wacker, S. A., Donow, C., Gessert, S., and Knchel, W. (2007) FoxN3 is required for craniofacial and eye development of *Xenopus laevis*. *Developmental Dynamics*, **236**(1), 226–239.
- [63] Samaan, G., Yugo, D., Rajagopalan, S., Wall, J., Donnell, R., Goldowitz, D., Gopalakrishnan, R., and Venkatachalam, S. (September, 2010) Foxn3 is essential for craniofacial development in mice and a putative candidate involved in human congenital craniofacial defects. *Biochemical and Biophysical Research Communications*, **400**(1), 60–65.
- [64] Brum, A. M., van de Peppel, J., van der Leije, C. S., Schreuders-Koedam, M., Eijken, M., van der Eerden, B. C. J., and van Leeuwen, J. P. T. M. (October, 2015) Connectivity Map-based discovery of parabendazole reveals targetable human osteogenic pathway. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(41), 12711–12716.

- [65] del Real, A., Prez-Campo, F. M., Fernndez, A. F., Saudo, C., Ibarbia, C. G., Prez-Nez, M. I., Crieckinge, W. V., Braspenning, M., Alonso, M. A., Fraga, M. F., and Riancho, J. A. (December, 2016) Differential analysis of genome-wide methylation and gene expression in mesenchymal stem cells of patients with fractures and osteoarthritis. *Epigenetics*, **12**(2), 113–122.
- [66] Bai, J., Yao, B., Wang, L., Sun, L., Chen, T., Liu, R., Yin, G., Xu, Q., and Yang, W. (2019) lncRNA A1BG-AS1 suppresses proliferation and invasion of hepatocellular carcinoma cells by targeting miR-216a-5p. *Journal of Cellular Biochemistry*, **120**(6), 10310–10322.
- [67] Li, N., Lee, W. Y.-W., Lin, S.-E., Ni, M., Zhang, T., Huang, X.-R., Lan, H.-Y., and Li, G. (October, 2014) Partial loss of Smad7 function impairs bone remodeling, osteogenesis and enhances osteoclastogenesis in mice. *Bone*, **67**, 46–55.
- [68] Vishal, M., Vimalraj, S., Ajeetha, R., Gokulnath, M., Keerthana, R., He, Z., Partridge, N. C., and Selvamurugan, N. (2017) MicroRNA-590-5p Stabilizes Runx2 by Targeting Smad7 During Osteoblast Differentiation. *Journal of Cellular Physiology*, **232**(2), 371–380.
- [69] Nowak, W. N., Taha, H., Kachamakova-Trojanowska, N., Stpniewski, J., Markiewicz, J. A., Kusienicka, A., Szade, K., Szade, A., Bukowska-Strakova, K., Hajduk, K., Klska, D., Kopacz, A., Grochot-Przczek, A., Barthenheier, K., Cauvin, C., Dulak, J., and Jzkowicz, A. (October, 2017) Murine Bone Marrow Mesenchymal Stromal Cells Respond Efficiently to Oxidative Stress Despite the Low Level of Heme Oxygenases 1 and 2. *Antioxidants & Redox Signaling*, **29**(2), 111–127.
- [70] Balogh, E., Paragh, G., and Jeney, V. (October, 2018) Influence of Iron on Bone Homeostasis. *Pharmaceuticals*, **11**(4).
- [71] Puri, N., Sodhi, K., Haarstad, M., Kim, D. H., Bohinc, S., Foglio, E., Favero, G., and Abraham, N. G. (June, 2012) Heme Induced Oxidative Stress Attenuates Sirtuin1 and Enhances Adipogenesis in Mesenchymal Stem Cells and Mouse Pre-Adipocytes. *Journal of Cellular Biochemistry*, **113**(6), 1926–1935.
- [72] Luo, Y., Tao, H., Jin, L., Xiang, W., and Guo, W. (November, 2019) CDKN2B-AS1 Exerts Oncogenic Role in Osteosarcoma by Promoting Cell Proliferation and Epithelial to Mesenchymal Transition. *Cancer Biotherapy and Radiopharmaceuticals*,.
- [73] Congrains, A., Kamide, K., Ohishi, M., and Rakugi, H. (January, 2013) ANRIL: Molecular Mechanisms and Implications in Human Health. *International Journal of Molecular Sciences*, **14**(1), 1278–1292.
- [74] Yin, Z., Ding, H., He, E., Chen, J., and Li, M. (October, 2016) Overexpression of long non-coding RNA MFI2 promotes cell proliferation and suppresses apoptosis in human osteosarcoma. *Oncology Reports*, **36**(4), 2033–2040.
- [75] Li, C., Tan, F., Pei, Q., Zhou, Z., Zhou, Y., Zhang, L., Wang, D., and Pei, H. (2019) Non-coding RNA MFI2-AS1 promotes colorectal cancer cell proliferation, migration and invasion through miR-574-5p/MYCBP axis. *Cell Proliferation*, **52**(4), e12632.

- [76] Zhu, C., Huang, L., Xu, F., Li, P., Li, P., and Hu, F. (October, 2019) LncRNA PCAT6 promotes tumor progression in osteosarcoma via activation of TGF- pathway by sponging miR-185-5p. *Biochemical and Biophysical Research Communications*,.
- [77] Dong, P., Xiong, Y., Yue, J., Hanley, S. J. B., Kobayashi, N., Todo, Y., and Watari, H. (October, 2018) Long Non-coding RNA NEAT1: A Novel Target for Diagnosis and Therapy in Human Tumors. *Frontiers in Genetics*, **9**.
- [78] Ahmed, A. S. I., Dong, K., Liu, J., Wen, T., Yu, L., Xu, F., Kang, X., Osman, I., Hu, G., Bunting, K. M., Crethers, D., Gao, H., Zhang, W., Liu, Y., Wen, K., Agarwal, G., Hirose, T., Nakagawa, S., Vazdarjanova, A., and Zhou, J. (September, 2018) Long noncoding RNA NEAT1 (nuclear paraspeckle assembly transcript 1) is critical for phenotypic switching of vascular smooth muscle cells. *Proceedings of the National Academy of Sciences*, **115**(37), E8660–E8667.
- [79] Taiana, E., Favasuli, V., Ronchetti, D., Todoerti, K., Pelizzoni, F., Manzoni, M., Barbieri, M., Fabris, S., Silvestris, I., Cantafio, M. E. G., Platonova, N., Zuccal, V., Maltese, L., Soncini, D., Ruberti, S., Cea, M., Chiaramonte, R., Amodio, N., Tassone, P., Agnelli, L., and Neri, A. (August, 2019) Long non-coding RNA NEAT1 targeting impairs the DNA repair machinery and triggers anti-tumor activity in multiple myeloma. *Leukemia*, pp. 1–11.
- [80] Wan, G., Mathur, R., Hu, X., Liu, Y., Zhang, X., Peng, G., and Lu, X. (May, 2013) Long non-coding RNA ANRIL (CDKN2B-AS) is induced by the ATM-E2F1 signaling pathway. *Cellular signalling*, **25**(5), 1086–1095.
- [81] Ding, K., Liao, Y., Gong, D., Zhao, X., and Ji, W. (July, 2018) Effect of long non-coding RNA H19 on oxidative stress and chemotherapy resistance of CD133+ cancer stem cells via the MAPK/ERK signaling pathway in hepatocellular carcinoma. *Biochemical and Biophysical Research Communications*, **502**(2), 194–201.
- [82] Yu, J.-L., Li, C., Che, L.-H., Zhao, Y.-H., and Guo, Y.-B. (2019) Downregulation of long noncoding RNA H19 rescues hippocampal neurons from apoptosis and oxidative stress by inhibiting IGF2 methylation in mice with streptozotocin-induced diabetes mellitus. *Journal of Cellular Physiology*, **234**(7), 10655–10670.
- [83] Hazell, G. G. J., Peachey, A. M. G., Teasdale, J. E., Sala-Newby, G. B., Angelini, G. D., Newby, A. C., and White, S. J. (December, 2016) PI16 is a shear stress and inflammation-regulated inhibitor of MMP2. *Scientific Reports*, **6**.
- [84] Puvvula, P. K. (May, 2019) LncRNAs Regulatory Networks in Cellular Senescence. *International Journal of Molecular Sciences*, **20**(11).
- [85] Spanner, M., Weber, K., Lanske, B., Ihbe, A., Siggelkow, H., Schtze, H., and Atkinson, M. J. (August, 1995) The iron-binding protein ferritin is expressed in cells of the osteoblastic lineage in vitro and in vivo. *Bone*, **17**(2), 161–165.
- [86] Balogh, E., Tolnai, E., Nagy, B., Nagy, B., Balla, G., Balla, J., and Jeney, V. (September, 2016) Iron overload inhibits osteogenic commitment and differentiation of mesenchymal stem cells via the induction of ferritin. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, **1862**(9), 1640–1649.

- [87] Zarjou, A., Jeney, V., Arosio, P., Poli, M., Antal-Szalms, P., Agarwal, A., Balla, G., and Balla, J. (June, 2009) Ferritin Prevents Calcification and Osteoblastic Differentiation of Vascular Smooth Muscle Cells. *Journal of the American Society of Nephrology*, **20**(6), 1254–1263.
- [88] Doi, M., Nagano, A., and Nakamura, Y. (January, 2002) Genome-wide Screening by cDNA Microarray of Genes Associated with Matrix Mineralization by Human Mesenchymal Stem Cells in Vitro. *Biochemical and Biophysical Research Communications*, **290**(1), 381–390.
- [89] Liu, Z., Zheng, Z., Qi, J., Wang, J., Zhou, Q., Hu, F., Liang, J., Li, C., Zhang, W., and Zhang, X. (December, 2018) CD24 identifies nucleus pulposus progenitors/notochordal cells for disc regeneration. *Journal of Biological Engineering*, **12**(1), 35.
- [90] Tsai, Y.-H., Lin, K.-L., Huang, Y.-P., Hsu, Y.-C., Chen, C.-H., Chen, Y., Sie, M.-H., Wang, G.-J., and Lee, M.-J. (2015) Suppression of ornithine decarboxylase promotes osteogenic differentiation of human bone marrow-derived mesenchymal stem cells. *FEBS Letters*, **589**(16), 2058–2065.
- [91] Chang, C.-F., Hsu, K.-H., Shen, C.-N., Li, C.-L., and Lu, J. (2014) Enrichment and Characterization of Two Subgroups of Committed Osteogenic Cells in the Mouse Endosteal Bone Marrow with Expression Levels of CD24. *Journal of Bone Research*, **2**(2), 1–9.
- [92] Park, G. C., Song, J. S., Park, H.-Y., Shin, S.-C., Jang, J. Y., Lee, J.-C., Wang, S.-G., Lee, B.-J., and Jung, J.-S. (May, 2016) Role of Fibroblast Growth Factor-5 on the Proliferation of Human Tonsil-Derived Mesenchymal Stem Cells. *Stem Cells and Development*, **25**(15), 1149–1160.
- [93] Kornmann, M., Ishiwata, T., Beger, H. G., and Korc, M. (September, 1997) Fibroblast growth factor-5 stimulates mitogenic signaling and is overexpressed in human pancreatic cancer: evidence for autocrine and paracrine actions. *Oncogene*, **15**(12), 1417–1424.
- [94] Williamson, E. A., Wray, J. W., Bansal, P., and Hromas, R. (2012) Overview for the Histone Codes for DNA Repair. *Progress in molecular biology and translational science*, **110**, 207–227.
- [95] Wang, S., Hu, B., Ding, Z., Dang, Y., Wu, J., Li, D., Liu, X., Xiao, B., Zhang, W., Ren, R., Lei, J., Hu, H., Chen, C., Chan, P., Li, D., Qu, J., Tang, F., and Liu, G.-H. (January, 2018) ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells. *Cell Discovery*, **4**.
- [96] Fu, L., Hu, Y., Song, M., Liu, Z., Zhang, W., Yu, F.-X., Wu, J., Wang, S., Izpisua Belmonte, J. C., Chan, P., Qu, J., Tang, F., and Liu, G.-H. (April, 2019) Up-regulation of FOXD1 by YAP alleviates senescence and osteoarthritis. *PLoS Biology*, **17**(4).
- [97] Samsonraj, R. M., Dudakovic, A., Manzar, B., Sen, B., Dietz, A. B., Cool, S. M., Rubin, J., and van Wijnen, A. J. (December, 2017) Osteogenic Stimulation of Human Adipose-Derived Mesenchymal Stem Cells Using a Fungal Metabolite That Suppresses the Polycomb Group Protein EZH2. *Stem Cells Translational Medicine*, **7**(2), 197–209.

- [98] Dudakovic, A., Gluscevic, M., Paradise, C. R., Dudakovic, H., Khani, F., Thaler, R., Ahmed, F. S., Li, X., Dietz, A. B., Stein, G. S., Montecino, M. A., Deyle, D. R., Westendorf, J. J., and van Wijnen, A. J. (April, 2017) Profiling of human epigenetic regulators using a semi-automated real-time qPCR platform validated by next generation sequencing. *Gene*, **609**, 28–37.
- [99] Camilleri, E. T., Gustafson, M. P., Dudakovic, A., Riester, S. M., Garces, C. G., Paradise, C. R., Takai, H., Karperien, M., Cool, S., Sampen, H.-J. I., Larson, A. N., Qu, W., Smith, J., Dietz, A. B., and van Wijnen, A. J. (August, 2016) Identification and validation of multiple cell surface markers of clinical-grade adipose-derived mesenchymal stromal cells as novel release criteria for good manufacturing practice-compliant production. *Stem Cell Research & Therapy*, **7**.
- [100] Knight, C., James, S., Kuntin, D., Fox, J., Newling, K., Hollings, S., Pennock, R., and Genever, P. (January, 2019) Epidermal growth factor can signal via -catenin to control proliferation of mesenchymal stem cells independently of canonical Wnt signalling. *Cellular Signalling*, **53**, 256–268.
- [101] Jiang, S., Cheng, S.-J., Ren, L.-C., Wang, Q., Kang, Y.-J., Ding, Y., Hou, M., Yang, X.-X., Lin, Y., Liang, N., and Gao, G. (September, 2019) An expanded landscape of human long noncoding RNA. *Nucleic Acids Research*, **47**(15), 7842–7856.
- [102] Chang, T.-H., Huang, H.-D., Ong, W.-K., Fu, Y.-J., Lee, O. K., Chien, S., and Ho, J. H. (April, 2014) The effects of actin cytoskeleton perturbation on keratin intermediate filament formation in mesenchymal stem/stromal cells. *Biomaterials*, **35**(13), 3934–3944.
- [103] Chang, Y., Li, H., and Guo, Z. (2014) Mesenchymal stem cell-like properties in fibroblasts. *Cellular Physiology and Biochemistry: International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology*, **34**(3), 703–714.
- [104] Denu, R. A., Nemcek, S., Bloom, D. D., Goodrich, A. D., Kim, J., Mosher, D. F., and Hematti, P. (August, 2016) Fibroblasts and Mesenchymal Stromal/Stem Cells Are Phenotypically Indistinguishable. *Acta Haematologica*, **136**(2), 85–97.
- [105] Ball, S. G., Shuttleworth, A. C., and Kielty, C. M. (April, 2004) Direct cell contact influences bone marrow mesenchymal stem cell fate. *The International Journal of Biochemistry & Cell Biology*, **36**(4), 714–727.
- [106] Tamama, K., Sen, C. K., and Wells, A. (October, 2008) Differentiation of Bone Marrow Mesenchymal Stem Cells into the Smooth Muscle Lineage by Blocking ERK/MAPK Signaling Pathway. *Stem Cells and Development*, **17**(5), 897–908.
- [107] Kumar, A., DSouza, S. S., Moskvina, O. V., Toh, H., Wang, B., Zhang, J., Swanson, S., Guo, L.-W., Thomson, J. A., and Slukvin, I. I. (May, 2017) Specification and Diversification of Pericytes and Smooth Muscle Cells from Mesenchymalangioblasts. *Cell Reports*, **19**(9), 1902–1916.

5 Figure Legends

5.1 Figure 1: Flowchart representation of the Pipeline used for this in the study.

The 4 steps of the flowchart are described A) Ab initio reconstruction of transcript expressed in MSC from SRA dataset and creation of a reference (gtf+fasta) for quantification of Ensembl annotated genes, unannotated intergenic (Mlincs) and unannotated overlapping antisens (Mloanc). The results are shown in Fig2.

B) Differential Analysis for the selection of MSC markers (restrained candidate set) with i/ kallisto pseudoalignment and Sleuth differential test followed by feature selection by random forest with Boruta package. Long read sequencing and active transcription in MSC by epigenetic marks information completed the selection step (see figures 2 and 3).

C) Validation of cell expression specificity of candidates by kmer quantification in ENCODE RNAseq datasets (see table S2 for list of data) and qPCR validation. The results are presented in figure 4.

D) Functional investigations were performed with in silico prediction methods from the sequence of candidates, followed by k-mer quantification with FANTOM6 data set, single cell RNAseq and selected MSC conditions. Kmer quantification phases are shown by corresponding icons (figures 5 and 6).

5.2 Figure 2: Overview of annotated genes and unannotated transcripts enriched in BM-MS.

A) Left pannels represented: i/EnsemblV90 transcript categories and distribution, ii/ transcripts distribution expressed in MSCs, showing unannotated transcripts obtained with Ab initio reconstruction by StringTie vs annotated transcripts (expression > 0.1 TPM) iii/ Predicted long non-coding RNA(lncRNA) from unannotated reconstructed transcripts include new long non coding RNA with intergenic (Mlinc) and antisens (Mlncoa) RNA categories.

B-C-D) Distribution of transcript length, exon length and GC percentage across different categories respectively with the same colors as in A pannel : coding transcripts (blue), annotated lincRNA (pink), annotated overlapping antisens lncRNA (purple), novel lincRNA (Mlincs, yellow), novel overlapping antisens RNA (Mloanc, red).

E) Representation of annotated genes (top pannel) and unannotated transcripts (bottom pannel) overexpressed in MSC versus non-MS types ($\log_2FC > 0.5$ and $p_{adj} < 0.05$), separately showed in MA plot.

F) Total number of transcripts transcript by category. The colored bar indicated the number of differential expression of annotated genes (Ensemblv90) and unannotated transcripts (Mlinc and Mloanc).

G) Global expression in BM-MS (with Sleuth normalization) of the same categories as in F for annotated genes and unannotated transcripts.

5.3 Figure 3: Selection of a refined set of best candidates by random forest (top35), long read sequencing and epigenetic features.

A) Expression of the best MSC-specific candidates selected by boruta machine learning along MSC group and not MSC cohorts. Left pannel : top35 most relevant annotated genes (non-coding included); Right pannel: unannotated intergenic lncRNAs (Mlincs) and their average importance scores determined by Boruta method displayed in upside line plot.

B) Genomic visualisation of Mlincs predictions 28428, 128022, 89912, and 64225 (MSClinc orange) from short reads alignment of all MSC group files (blue/magenta and bam visualisation), compared with long read alignments (long reads grey). Additional epigenomic features are shown to reveal active transcriptional activity from trimethylation of Histone H3 (H3K4me3), acetylation of Histone H3 H3K27 in MSCs (H3K4me3 and H3K27ac, green), and Dnase sensibility hotspots of MSC (MSC DNase, red).

5.4 Figure 4: High throughput exploration of selected candidate across strong variety of samples by k-mer quantification in RNAseq and biological validation by RT-qPCR.

A : List of tissues for the cell specific expression exploration (samples with ID numbers are listed in table S1)

B: Relative expression of Mlinc.28428.2, Mlinc.128022.2, and Mlinc.89912.1 across ENCODEs ribodepleted RNAseq datas, made by k-mer quantification, normalized by k-mer by million.

C: qPCR relative quantification was performed on the selected 3 Mlincs in MSC of different origins (BM-MSc, Ad-MSc, Umbilical cord msc) and other indicated cell types. Relative quantification (Log induction) was quantified by ddCt method using non MSC types types as calibrator (mean of triplicates). Student tests have been made between triplicates, each test use BM-MSCs as reference group (ns : $P > 0.05$, *: $P \leq 0.05$, **: $P \leq 0.01$, ***: $P \leq 0.001$, ****: $P \leq 0.0001$).

5.5 Figure 5: Prediction of potential functions of candidates with k-mer quantification and single-cell.

For each Mlinc (Mlinc.28428 (A), Mlinc.128022 (B) and Mlinc.89912 (C) respectively) 3 steps of prediction were performed. a. Enrichment in the different subcompartments of fibroblasts from FANTOM6 dataset: Free nuclear fraction (Nuc), chromatin (Chr) and cytoplasm (Cyt); b. Expression of marker in FANTOM6 data depending of the KnockDown (KD) of an annotated lncRNA. Normalised count of all specific k-mers is averaged by sample (zeros vlues deleted) and t-tests are made between control and KD fibroblasts. c. General expression of Mlincs inside Ad-MSc population, dimensionnal reduction made with UMAP method, made from batch corrected counts. Expression of differentially expressed annotated genes between positive and negative cells for Mlinc.28428, Mlinc.128022 and Mlinc.89912 respectively.

5.6 Figure 6: Expression of markers in different datasets from SRA in cell conditions related to previous findings

A) Expression of Minc.28428.1 in the context of oxydative, replicative, or KO-driven, stress and senescence (PRJNA396193, PRJNA433339). Relevant changes of expressions are showed with t-test results (ns : $P > 0.05$, *: $P \leq 0.05$, **: $P \leq 0.01$, ***: $P \leq 0.001$, ****: $P \leq 0.0001$).

B) Expression of Minc.128022 in osteodifferentiation conditions (PRJNA515466) or osteodifferentiation potential (PRJNA379707). Relevant changes of expressions are showed with t-test results (ns : $P > 0.05$, *: $P \leq 0.05$, **: $P \leq 0.01$, ***: $P \leq 0.001$, ****: $P \leq 0.0001$).

C) Expression of Minc.89912 in the context of proliferation (PRJNA328824 and PRJNA498109). Relevant changes of expressions are showed with t-test results (ns : $P > 0.05$, *: $P \leq 0.05$, **: $P \leq 0.01$, ***: $P \leq 0.001$, ****: $P \leq 0.0001$). The detailed list of datasets is provided in table S4.

Figure 1

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.09.976001>; this version posted March 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

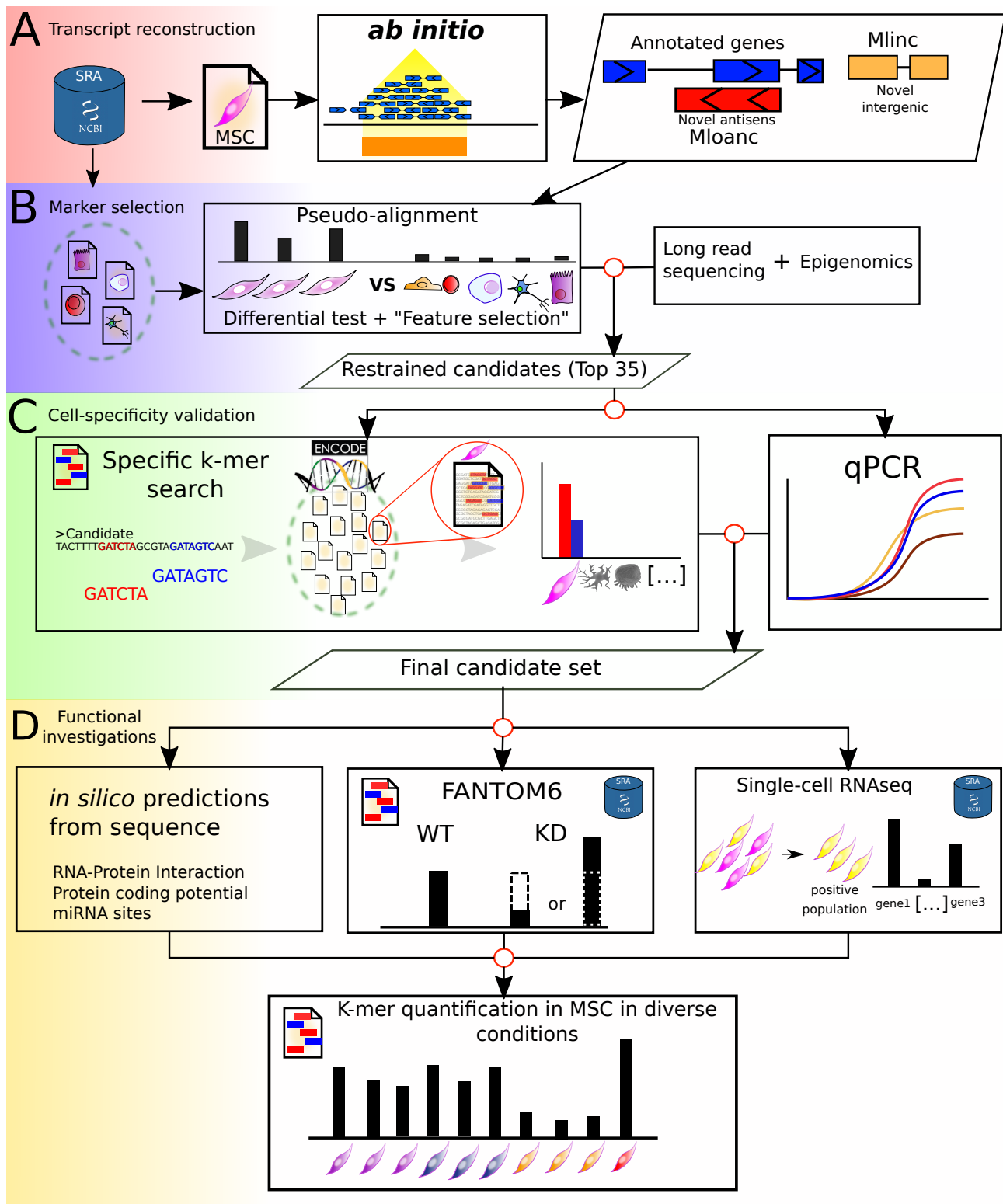
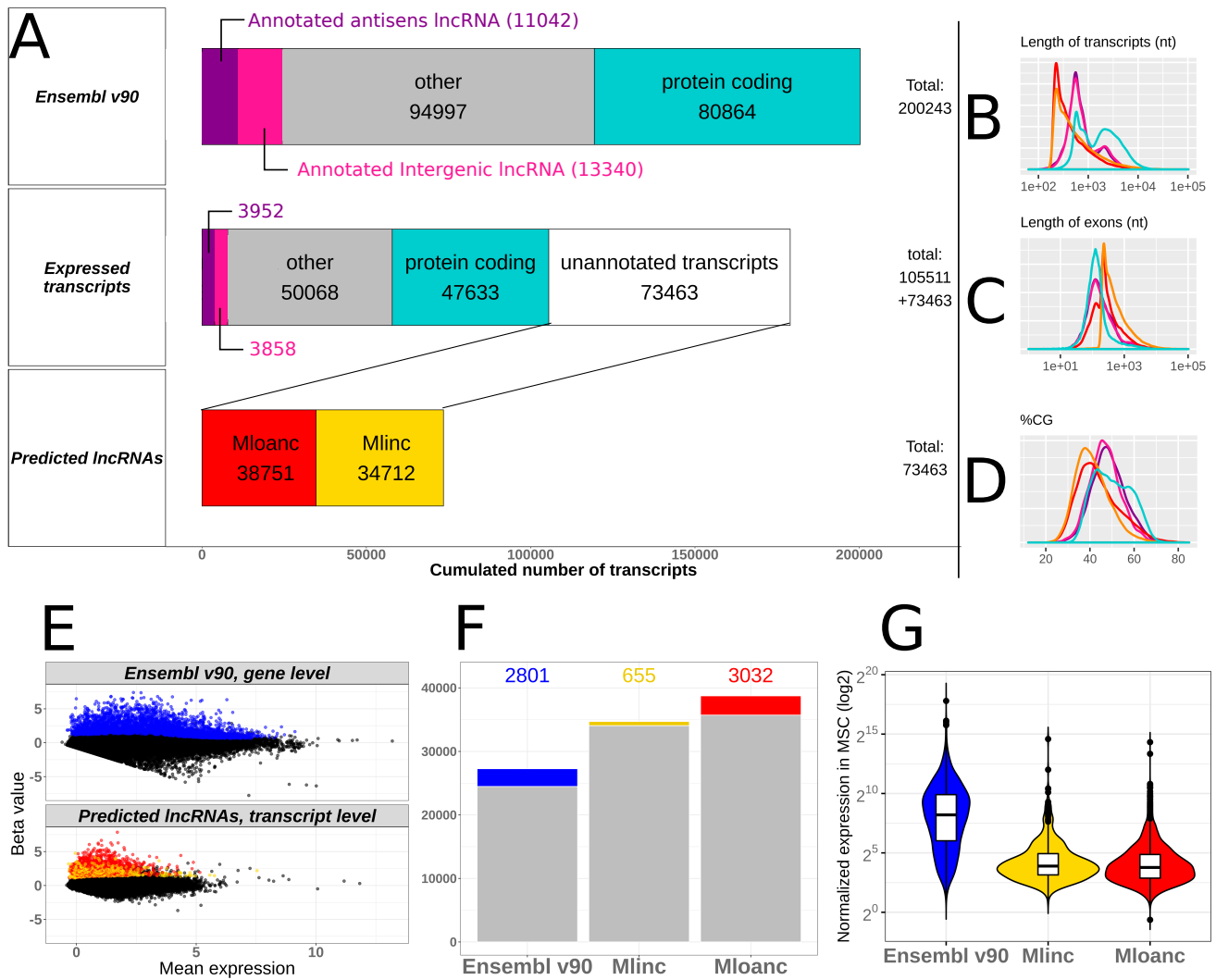


Figure 2

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.09.976001>; this version posted March 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.



bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.09.976001>; this version posted March 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

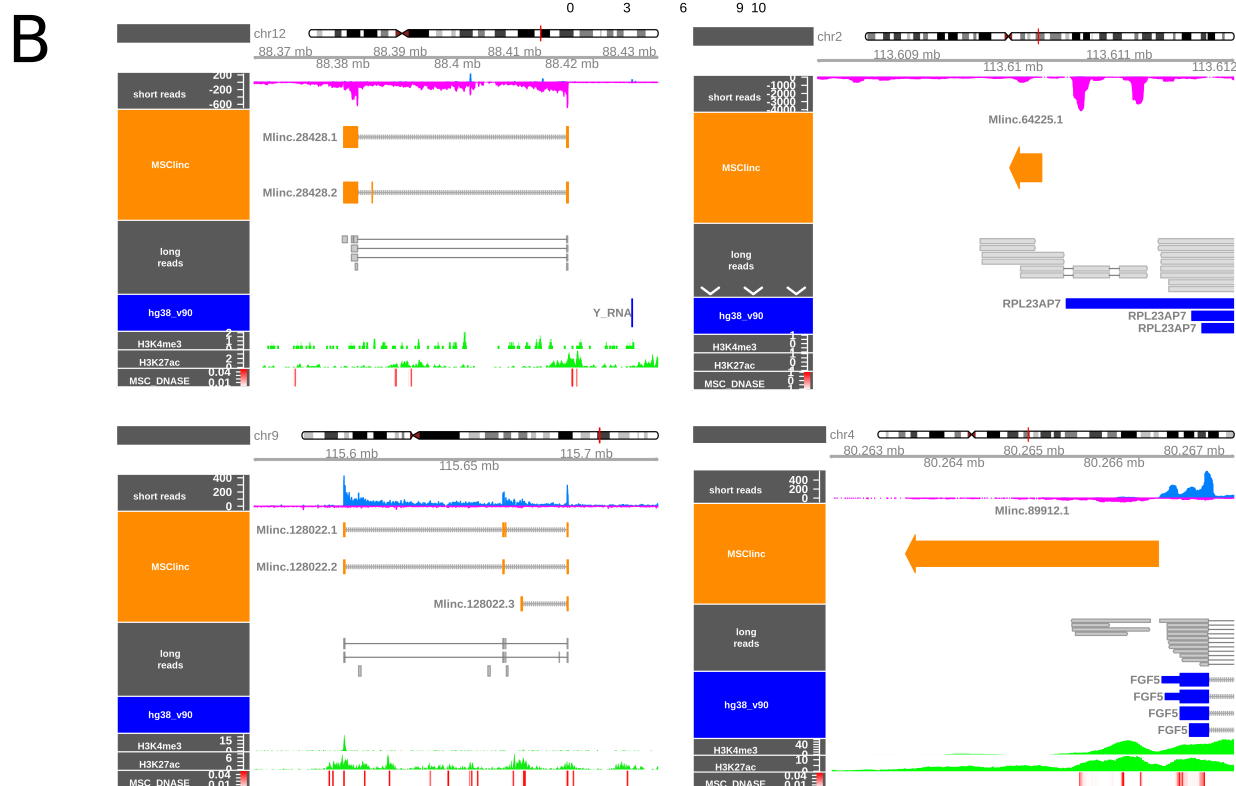
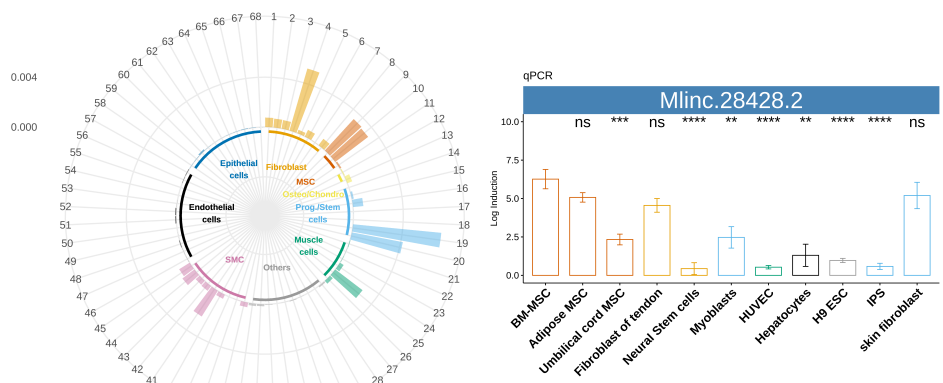


Figure 4

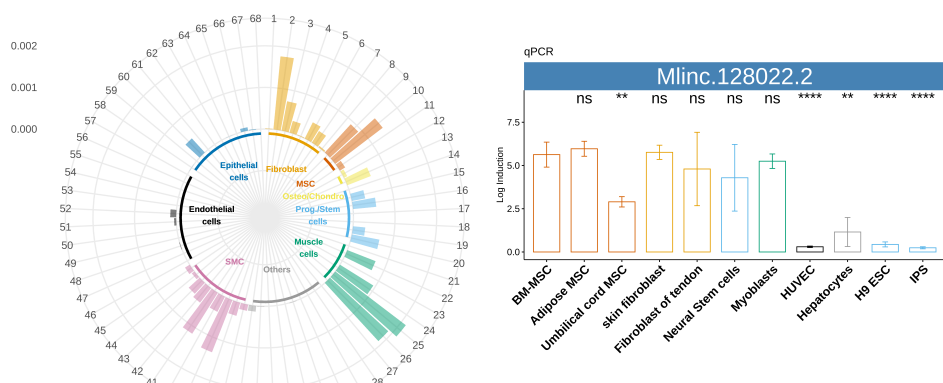
bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.09.976001>; this version posted March 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

id	variable	n
1	bronchus fibroblast of lung	2
2	cardiac atrium fibroblast	2
3	cardiac ventricle fibroblast	2
4	fibroblast of dermis	2
5	fibroblast of lung	2
6	fibroblast of the aortic adventitia	2
7	fibroblast of villous mesenchyme	2
8	pericardium fibroblast	2
9	MSC of adipose	2
10	MSC of the bone marrow	2
11	MSC of Wharton's jelly	2
12	articular chondrocyte	2
13	osteoblast	2
14	H1-hESC	5
15	H7-hESC	2
16	hair follicle dermal papilla cell	2
17	hematopoietic multipotent progenitor cell	1
18	neural progenitor cell	2
19	skeletal muscle satellite cell	2
20	subcutaneous preadipocyte	2
21	cardiac muscle cell	2
22	myocyte	2
23	myometrial cell	2
24	myotube	2
25	regular cardiac myocyte	2
26	skeletal muscle myoblast	2
27	astrocyte	2
28	bipolar neuron	2
29	dendritic cell	20
30	foreskin keratinocyte	6
31	hair follicular keratinocyte	2
32	hepatocyte	2
33	melanocyte of skin	4
34	mesangial cell	2
35	mononuclear cell	1
36	placental pericyte	2
37	aortic SMC	2
38	bronchial smooth muscle cell	2
39	smooth muscle cell	2
40	smooth muscle cell of bladder	2
41	smooth muscle cell of the coronary artery	2
42	smooth muscle cell of the pulmonary artery	2
43	smooth muscle cell of the umbilical artery	2
44	smooth muscle cell of trachea	2
45	uterine smooth muscle cell	2
46	bladder microvascular endothelial cell	2
47	dermis blood vessel endothelial cell	2
48	dermis lymphatic vessel endothelial cell	2
49	dermis microvascular lymphatic vessel endothelial cell	2
50	endometrial microvascular endothelial cells	2
51	endothelial cell of coronary artery	2
52	glomerular endothelial cell	2
53	lung microvascular endothelial cell	2
54	mammary microvascular endothelial cell	2
55	pulmonary artery endothelial cell	2
56	thoracic aorta endothelial cell	2
57	vein endothelial cell	2
58	airway epithelial cell	2
59	bronchial epithelial cell	2
60	epithelial cell of alveolus of lung	2
61	epithelial cell of proximal tubule	2
62	artery epithelial cell of umbilical	2
63	kidney epithelial cell	2
64	mammary epithelial cell	1
65	epithelial cell of viscerocranial mucosa	2
66	placental epithelial cell	2
67	renal cortical epithelial cell	2
68	tracheal epithelial cell	2

Mlinc.28428.2



Mlinc.128022.2



Mlinc.89912.1

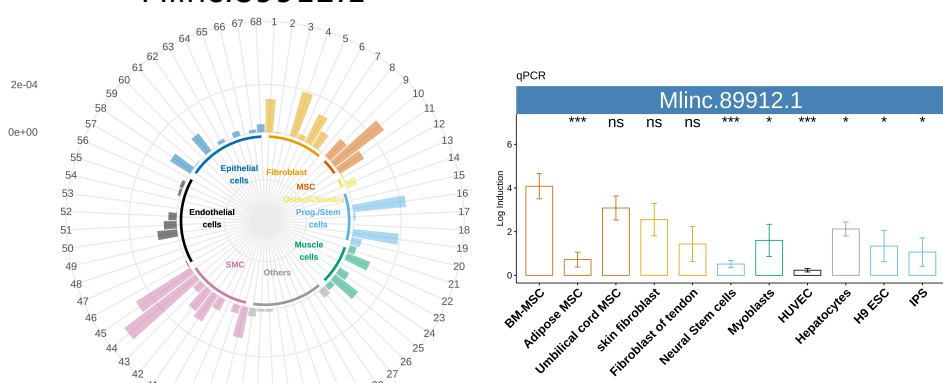
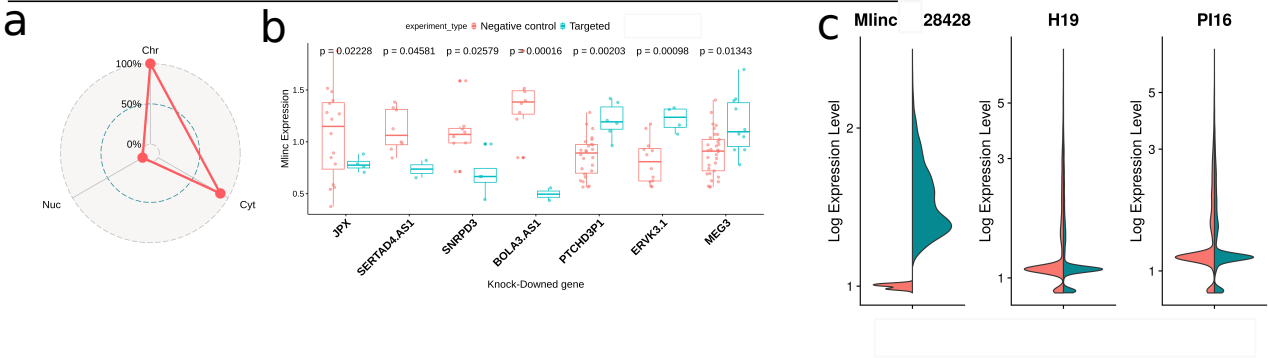


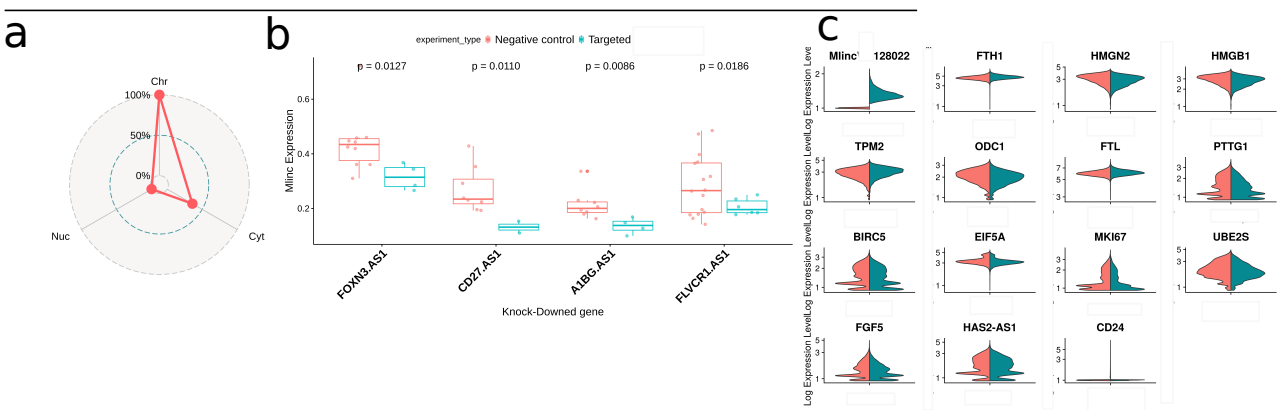
Figure 5

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.09.976001>; this version posted March 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

A: Mlinc.28428



B: Mlinc.128022



C : Mlinc.89912

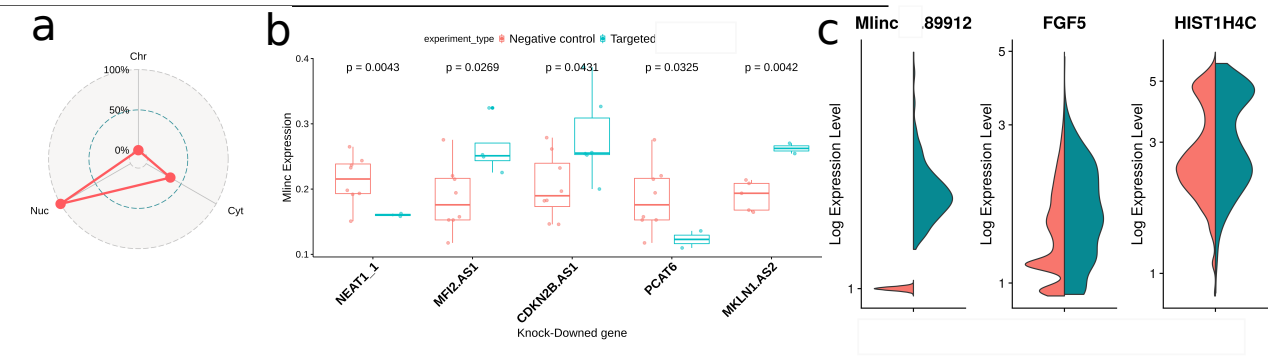
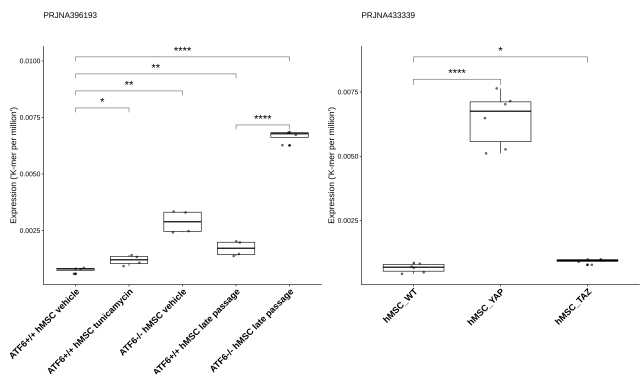


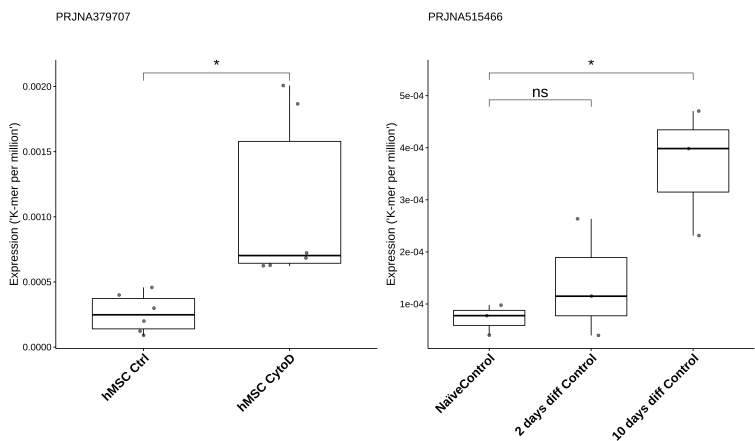
Figure 6

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.09.976001>; this version posted March 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

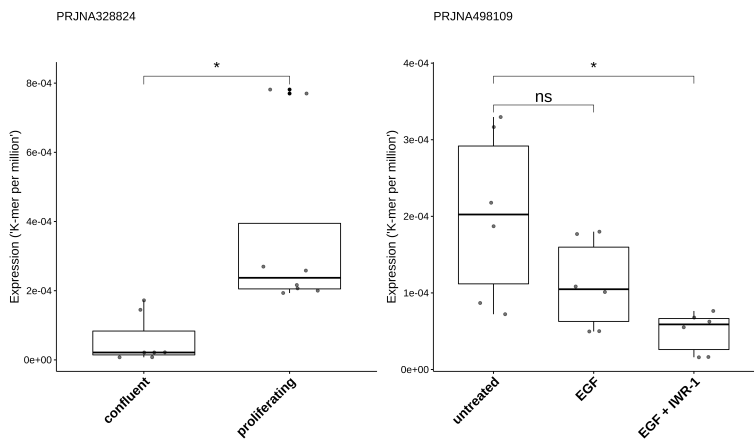
A: Mlinc.28428



B: Mlinc.128022



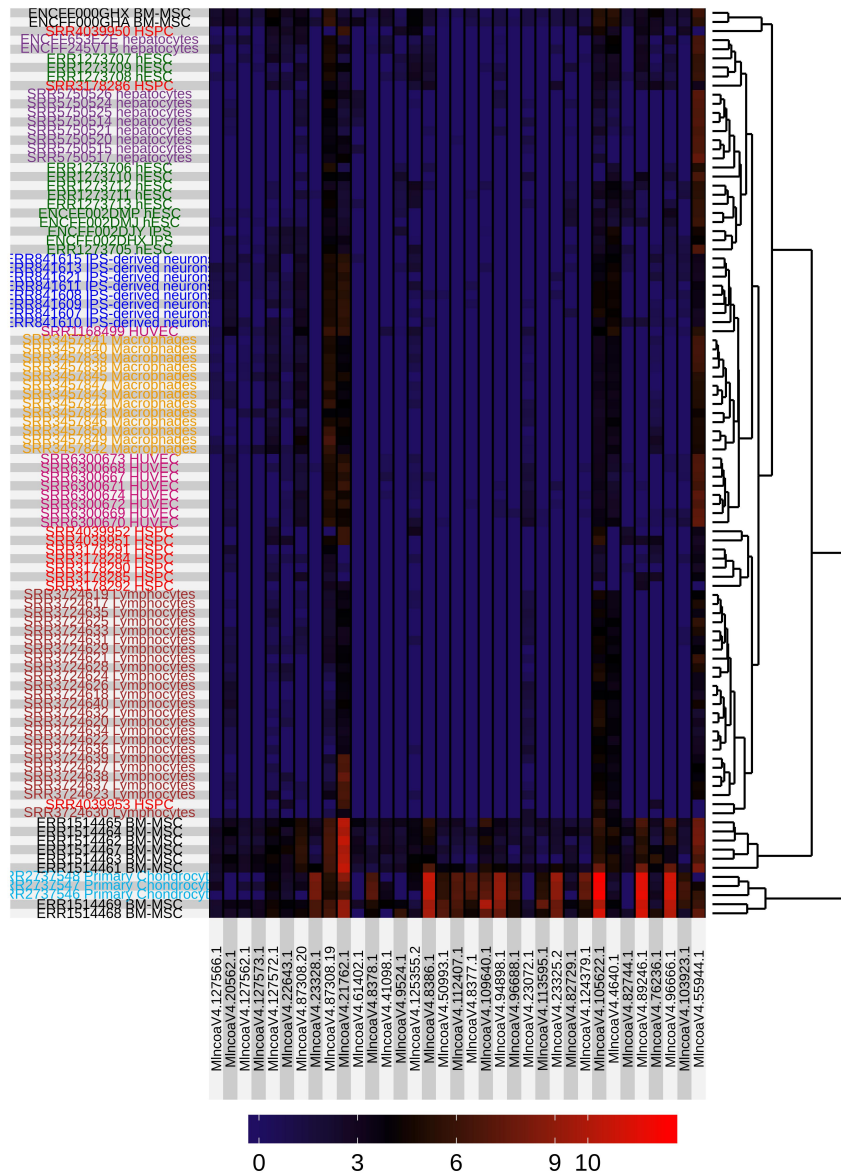
C : Mlinc.89912



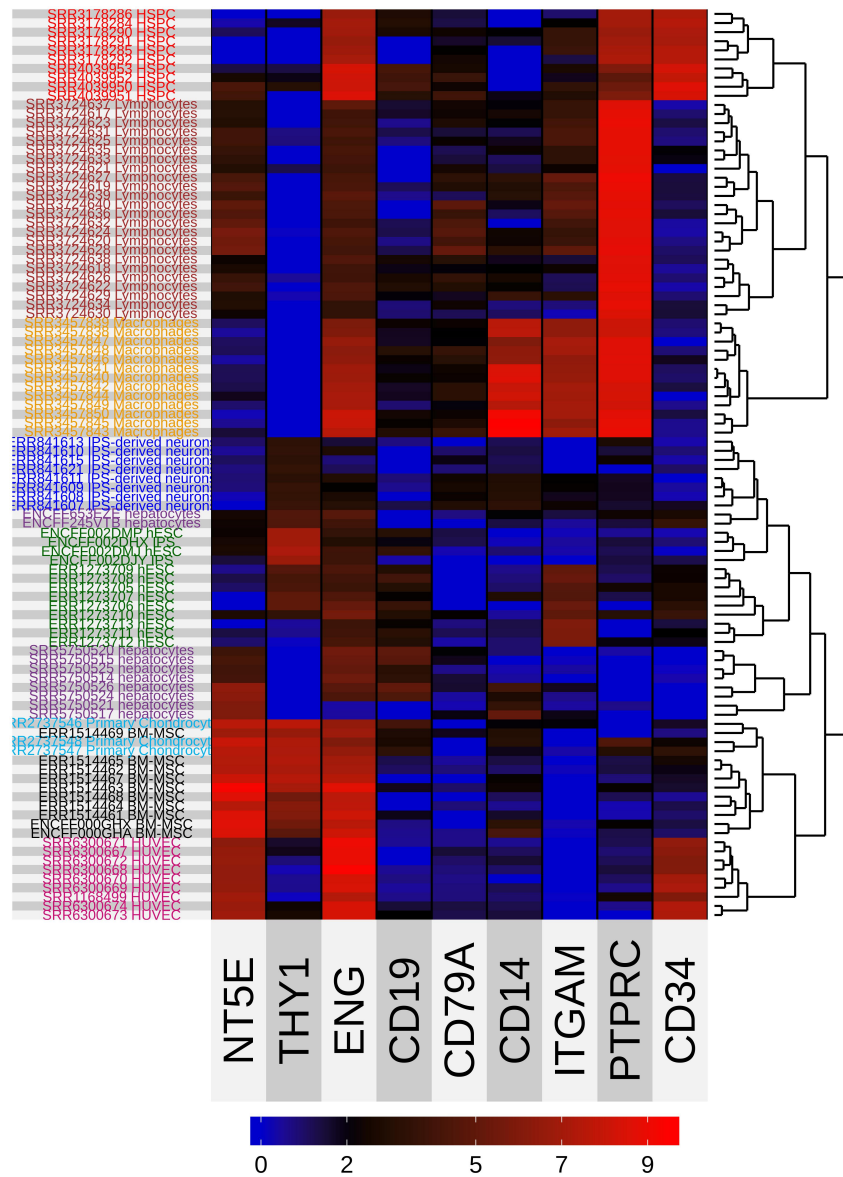
Supplementary figures

sebastien.riquier

January 2020



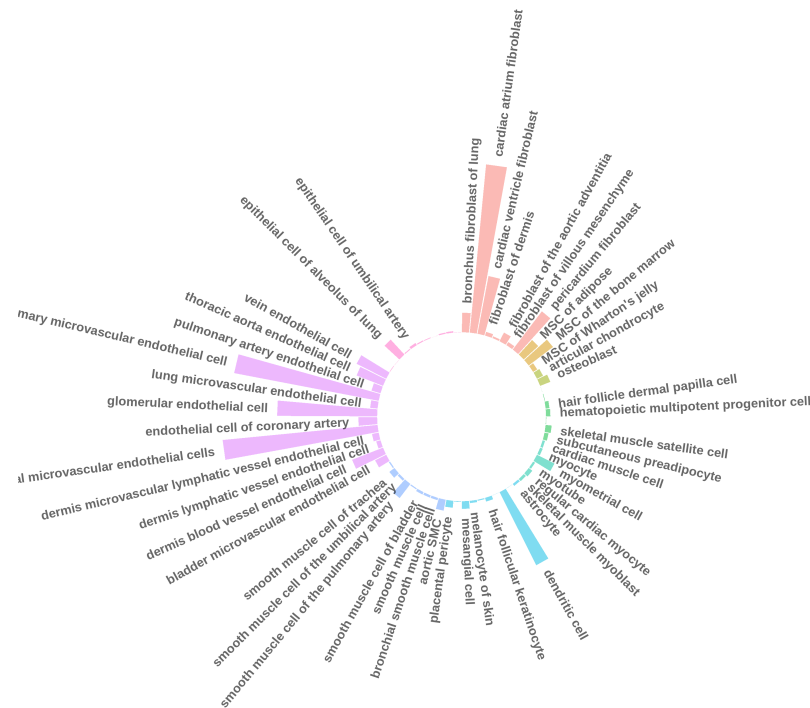
Supplementary Figure 1: Expression of Mloanc (Antisens unannotated RNA) selected after feature selection in the differential analysis cohort



Supplementary Figure 2: Expression of ISCT's MSC markers in the differential analysis cohort THY1 = CD90, NT5E = CD73, ENG = CD105, ITGAM = CD11B, PTPRC = CD45

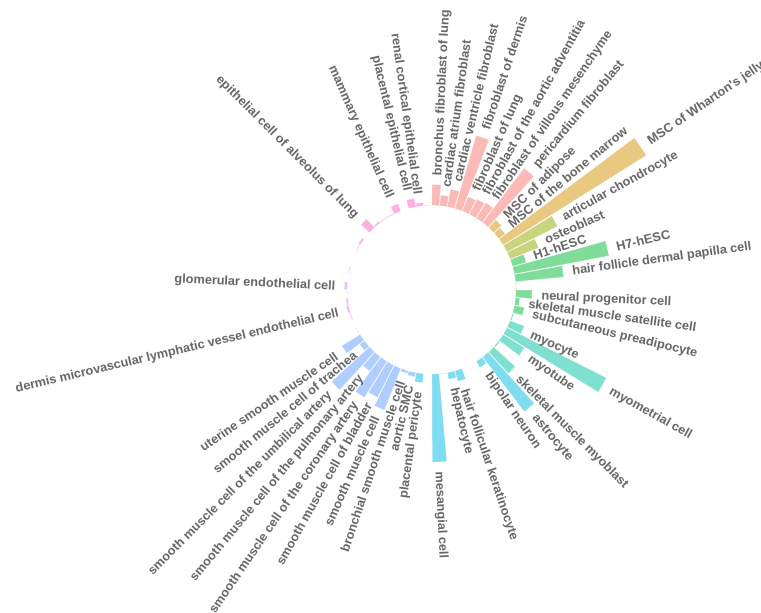
4

ENG-ENST00000344849

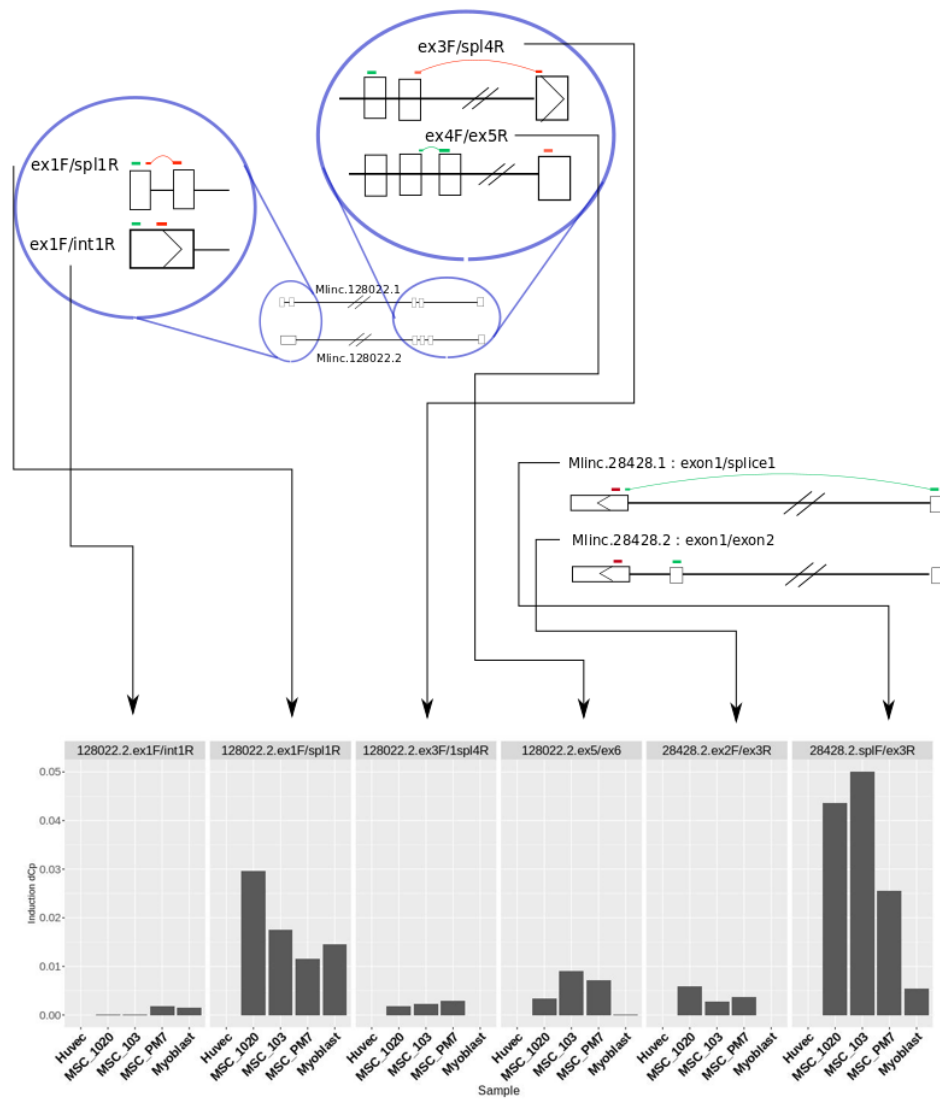


Supplementary Figure 4: Relative expression of 3 positive markers of ENG (CD105) across ENCODE's ribodepleted RNAseq datas, made by K-mer quantification, normalized in kmer by million

THY1-ENST00000284240



Supplementary Figure 6: Relative expression of 3 positive markers of THY1 (CD90) across ENCODE ribodepleted RNAseq datas, made by K-mer quantification, normalized in kmer by million



Supplementary Figure 8: Primer position on selected Mlinc candidates and corresponding expression in MSCs, HUVECs and Myoblasts



Supplementary Figure 9: cell cycle and single cell.

Table_S1 Metadata for differential analysis

sample ID	Group	Type	SRX ID	external ID	sra sample title	specie	project title	project ID	nb total reads	length reads	detail
ENCFF000GHA	MSC	BM-MSC			total RNA-seqof mesenchymal stem cell of the bone marrow	Homo sapiens	ENCODE				
ENCFF000GHX	MSC	BM-MSC			total RNA-seqof mesenchymal stem cell of the bone marrow	Homo sapiens	ENCODE				
ENCFF002DHX	Not	IPS			total RNA-seqof GM23338	Homo sapiens	ENCODE				
ENCFF002DJY	Not	IPS			total RNA-seqof GM23338	Homo sapiens	ENCODE				
ENCFF002DMJ	Not	hESC			total RNA-seqof H7	Homo sapiens	ENCODE				
ENCFF002DMP	Not	hESC			total RNA-seqof H7	Homo sapiens	ENCODE				
ERR1273705	Not	hESC	ERX1345312		Transcriptional profiling of human naive embryonic stem cells	Homo sapiens	Transcriptional profiling of human naive embryonic stem cells	PRJEB12748	67286349	125	
ERR1273706	Not	hESC	ERX1345313		Transcriptional profiling of human naive embryonic stem cells	Homo sapiens	Transcriptional profiling of human naive embryonic stem cells	PRJEB12748	18924239	125	
ERR1273707	Not	hESC	ERX1345314		Transcriptional profiling of human naive embryonic stem cells	Homo sapiens	Transcriptional profiling of human naive embryonic stem cells	PRJEB12748	55745935	125	
ERR1273708	Not	hESC	ERX1345315		Transcriptional profiling of human naive embryonic stem cells	Homo sapiens	Transcriptional profiling of human naive embryonic stem cells	PRJEB12748	28988509	125	
ERR1273709	Not	hESC	ERX1345316		Transcriptional profiling of human naive embryonic stem cells	Homo sapiens	Transcriptional profiling of human naive embryonic stem cells	PRJEB12748	NA	NA	
ERR1273710	Not	hESC	ERX1345317		Transcriptional profiling of human naive embryonic stem cells	Homo sapiens	Transcriptional profiling of human naive embryonic stem cells	PRJEB12748	NA	NA	
ERR1273711	Not	hESC	ERX1345318		Transcriptional profiling of human naive embryonic stem cells	Homo sapiens	Transcriptional profiling of human naive embryonic stem cells	PRJEB12748	NA	NA	
ERR1273712	Not	hESC	ERX1345319		Transcriptional profiling of human naive embryonic stem cells	Homo sapiens	Transcriptional profiling of human naive embryonic stem cells	PRJEB12748	NA	NA	
ERR1273713	Not	hESC	ERX1345320		Transcriptional profiling of human naive embryonic stem cells	Homo sapiens	Transcriptional profiling of human naive embryonic stem cells	PRJEB12748	NA	NA	
ERR1514461	MSC	BM-MSC	ERX1585480		RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	Homo sapiens	RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	PRJEB14688	32467641	96.41	
ERR1514462	MSC	BM-MSC	ERX1585481		RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	Homo sapiens	RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	PRJEB14688	34192962	96.30	
ERR1514463	MSC	BM-MSC	ERX1585482		RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	Homo sapiens	RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	PRJEB14688	28164147	95.51	
ERR1514464	MSC	BM-MSC	ERX1585483		RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	Homo sapiens	RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	PRJEB14688	28633110	96.97	
ERR1514465	MSC	BM-MSC	ERX1585484		RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	Homo sapiens	RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	PRJEB14688	25135660	96.96	
ERR1514467	MSC	BM-MSC	ERX1585486		RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	Homo sapiens	RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	PRJEB14688	28240410	96.79	
ERR1514468	MSC	BM-MSC	ERX1585487		RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	Homo sapiens	RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	PRJEB14688	29410953	96.97	
ERR1514469	MSC	BM-MSC	ERX1585488		RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	Homo sapiens	RNASeq of musculoskeletal ageing using human mesenchymal stem cells and their tissue constructs	PRJEB14688	28966063	97.22	
ERR841607	Not	IPS-derived neurons	ERX921697		illumina HiSeq 2500 paired end sequencing; RNA-seq of NEAT1 knockdown in neuronal excitability	Homo sapiens	RNA-seq of NEAT1 knockdown in neuronal excitability	PRJEB9006	14282989	101	KCI
ERR841608	Not	IPS-derived neurons	ERX921709		illumina HiSeq 2500 paired end sequencing; RNA-seq of NEAT1 knockdown in neuronal excitability	Homo sapiens	RNA-seq of NEAT1 knockdown in neuronal excitability	PRJEB9006	22188866	101	KCI
ERR841609	Not	IPS-derived neurons	ERX921694		illumina HiSeq 2500 paired end sequencing; RNA-seq of NEAT1 knockdown in neuronal excitability	Homo sapiens	RNA-seq of NEAT1 knockdown in neuronal excitability	PRJEB9006	22616373	101	KCI
ERR841610	Not	IPS-derived neurons	ERX921708		illumina HiSeq 2500 paired end sequencing; RNA-seq of NEAT1 knockdown in neuronal excitability	Homo sapiens	RNA-seq of NEAT1 knockdown in neuronal excitability	PRJEB9006	13192833	101	NA
ERR841611	Not	IPS-derived neurons	ERX921705		illumina HiSeq 2500 paired end sequencing; RNA-seq of NEAT1 knockdown in neuronal excitability	Homo sapiens	RNA-seq of NEAT1 knockdown in neuronal excitability	PRJEB9006	15290840	101	KCI
ERR841613	Not	IPS-derived neurons	ERX921699		illumina HiSeq 2500 paired end sequencing; RNA-seq of NEAT1 knockdown in neuronal excitability	Homo sapiens	RNA-seq of NEAT1 knockdown in neuronal excitability	PRJEB9006	15686003	101	NA
ERR841615	Not	IPS-derived neurons	ERX921702		illumina HiSeq 2500 paired end sequencing; RNA-seq of NEAT1 knockdown in neuronal excitability	Homo sapiens	RNA-seq of NEAT1 knockdown in neuronal excitability	PRJEB9006	14182869	101	NA
ERR841621	Not	IPS-derived neurons	ERX921696		illumina HiSeq 2500 paired end sequencing; RNA-seq of NEAT1 knockdown in neuronal excitability	Homo sapiens	RNA-seq of NEAT1 knockdown in neuronal excitability	PRJEB9006	15232542	101	NA
SRR2737546	Not	Primary Chondrocytes	SRX1358410		GSM1914777: Control 1; Homo sapiens; RNA-Seq	Homo sapiens	Transcriptome response to 4h IL-1b stimulation of primary chondrocytes	PRJNA299568	34300359	97.39	No stimulation
SRR2737547	Not	Primary Chondrocytes	SRX1358411		GSM1914778: Control 2; Homo sapiens; RNA-Seq	Homo sapiens	Transcriptome response to 4h IL-1b stimulation of primary chondrocytes	PRJNA299568	37284758	96.91	No stimulation
SRR2737548	Not	Primary Chondrocytes	SRX1358412		GSM1914779: Control 3; Homo sapiens; RNA-Seq	Homo sapiens	Transcriptome response to 4h IL-1b stimulation of primary chondrocytes	PRJNA299568	33144961	96.89	No stimulation
SRR4039950	Not	HSPC	SRX2031066		GSM2284651: HSPC day 0 rep1; Homo sapiens; RNA-Seq	Homo sapiens	Hematopoietic Stem/Progenitor Cell Expansion Without Differentiation is Achieved in Zwitterionic Hydrogels	PRJNA339416	17369972	50	
SRR4039951	Not	HSPC	SRX2031067		GSM2284652: HSPC day 0 rep2; Homo sapiens; RNA-Seq	Homo sapiens	Hematopoietic Stem/Progenitor Cell Expansion Without Differentiation is Achieved in Zwitterionic Hydrogels	PRJNA339416	18053125	50	
SRR4039952	Not	HSPC	SRX2031068		GSM2284653: HSPC day 24 rep1; Homo sapiens; RNA-Seq	Homo sapiens	Hematopoietic Stem/Progenitor Cell Expansion Without Differentiation is Achieved in Zwitterionic Hydrogels	PRJNA339416	20682268	50	
SRR4039953	Not	HSPC	SRX2031069		GSM2284654: HSPC day 24 rep2; Homo sapiens; RNA-Seq	Homo sapiens	Hematopoietic Stem/Progenitor Cell Expansion Without Differentiation is Achieved in Zwitterionic Hydrogels	PRJNA339416	19282208	50	
SRR3178284	Not	HSPC	SRX1592574		GSM2066393: huCB stem ctrl_rep1; Homo sapiens; RNA-Seq	Homo sapiens	miR-126 Orchestrates an Oncogenic Program in B-Cell Precursor Acute Lymphoblastic Leukemia	PRJNA312538	25398470	100.79	huCB stem ctrl
SRR3178285	Not	HSPC	SRX1592575		GSM2066394: huCB stem ctrl_rep2; Homo sapiens; RNA-Seq	Homo sapiens	miR-126 Orchestrates an Oncogenic Program in B-Cell Precursor Acute Lymphoblastic Leukemia	PRJNA312538	5926147	100.77	huCB stem ctrl
SRR3178286	Not	HSPC	SRX1592576		GSM2066395: huCB stem ctrl_rep3; Homo sapiens; RNA-Seq	Homo sapiens	miR-126 Orchestrates an Oncogenic Program in B-Cell Precursor Acute Lymphoblastic Leukemia	PRJNA312538	7615479	100.78	huCB stem ctrl
SRR3178290	Not	HSPC	SRX1592580		GSM2066399: huCB progenitors ctrl_rep1; Homo sapiens; RNA-Seq	Homo sapiens	miR-126 Orchestrates an Oncogenic Program in B-Cell Precursor Acute Lymphoblastic Leukemia	PRJNA312538	36341406	100.77	huCB progenitors ctrl
SRR3178291	Not	HSPC	SRX1592581		GSM2066400: huCB progenitors ctrl_rep2; Homo sapiens; RNA-Seq	Homo sapiens	miR-126 Orchestrates an Oncogenic Program in B-Cell Precursor Acute Lymphoblastic Leukemia	PRJNA312538	6431588	100.78	huCB progenitors ctrl
SRR3178292	Not	HSPC	SRX1592582		GSM2066401: huCB progenitors ctrl_rep3; Homo sapiens; RNA-Seq	Homo sapiens	miR-126 Orchestrates an Oncogenic Program in B-Cell Precursor Acute Lymphoblastic Leukemia	PRJNA312538	3602619	100.77	huCB progenitors ctrl
ENCFF245VTB	Not	hepatocytes			total RNA-seqof hepatocyte	Homo sapiens	ENCODE				
ENCFF653EZE	Not	hepatocytes			total RNA-seqof hepatocyte	Homo sapiens	ENCODE				
SRR1168499	Not	HUVEC	SRX470246		GSM1327344: HUVEC; Homo sapiens; RNA-Seq	Homo sapiens	Transcriptomic analysis reveals novel long non-coding RNAs critical for vertebrate development [RNA-Seq]	PRJNA238180	97606136	90	
SRR3457838	Not	Macrophages	SRX1733620		GSM2135456: 140421_S2_CNT_A_ ; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	62681472	50	CNT
SRR3457839	Not	Macrophages	SRX1733621		GSM2135457: 140421_S2_CNT_B_ ; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	66405838	50	CNT
SRR3457840	Not	Macrophages	SRX1733622		GSM2135458: 140421_S2_IL10_A_ ; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	70016991	50	IL10
SRR3457841	Not	Macrophages	SRX1733623		GSM2135459: 140421_S2_IL10_B_ ; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	80257003	50	IL10
SRR3457842	Not	Macrophages	SRX1733624		GSM2135460: 140825_S3_CNT_ ; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	37671849	75	CNT

Table_S1 Metadata for differential analysis

SRR3457843	Not	Macrophages	SRX1733625		GSM2135461: 140825_S3_IL10____; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	40575802	75	IL10	NA
SRR3457844	Not	Macrophages	SRX1733626		GSM2135462: 140825_S1_CNT____; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	28548046	75	CNT	NA
SRR3457845	Not	Macrophages	SRX1733627		GSM2135463: 140825_S1_IL10____; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	44848576	75	IL10	NA
SRR3457846	Not	Macrophages	SRX1733628		GSM2135464: 140825_S1_IFNgTNF____; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	33920601	75	IFNgTNF	NA
SRR3457847	Not	Macrophages	SRX1733629		GSM2135465: 140825_S2_IFNgTNF____; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	26580763	75	IFNgTNF	NA
SRR3457848	Not	Macrophages	SRX1733630		GSM2135466: 140825_S3_IFNgTNF____; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	28565135	75	IFNgTNF	NA
SRR3457849	Not	Macrophages	SRX1733631		GSM2135467: 140825_S2_CNT____; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	41595399	75	CNT	NA
SRR3457850	Not	Macrophages	SRX1733632		GSM2135468: 140825_S2_IL10____; Homo sapiens; RNA-Seq	Homo sapiens	Single-cell qPCR, and bulk-level transcriptomics and epigenomics of treated macrophages	PRJNA315538	38443748	75	IL10	NA
SRR5750514	Not	hepatocytes	SRX2950738		Dual RNAseq: Human hepatocytes and Plasmodium berghei	Homo sapiens	Dual RNAseq - Human hepatocytes infected with Plasmodium berghei	PRJNA390648	62983452	99.38	Infected with Plasmodium berghei	NA
SRR5750515	Not	hepatocytes	SRX2950737		Dual RNAseq: Human hepatocytes and Plasmodium berghei	Homo sapiens	Dual RNAseq - Human hepatocytes infected with Plasmodium berghei	PRJNA390648	46318597	99.38	NA	NA
SRR5750517	Not	hepatocytes	SRX2950735		Dual RNAseq: Human hepatocytes and Plasmodium berghei	Homo sapiens	Dual RNAseq - Human hepatocytes infected with Plasmodium berghei	PRJNA390648	63391464	99.29	NA	NA
SRR5750520	Not	hepatocytes	SRX2950732		Dual RNAseq: Human hepatocytes and Plasmodium berghei	Homo sapiens	Dual RNAseq - Human hepatocytes infected with Plasmodium berghei	PRJNA390648	57010988	99.33	NA	NA
SRR5750521	Not	hepatocytes	SRX2950731		Dual RNAseq: Human hepatocytes and Plasmodium berghei	Homo sapiens	Dual RNAseq - Human hepatocytes infected with Plasmodium berghei	PRJNA390648	47651456	99.41	Infected with Plasmodium berghei	NA
SRR5750524	Not	hepatocytes	SRX2950728		Dual RNAseq: Human hepatocytes and Plasmodium berghei	Homo sapiens	Dual RNAseq - Human hepatocytes infected with Plasmodium berghei	PRJNA390648	47486674	99.53	NA	NA
SRR5750525	Not	hepatocytes	SRX2950727		Dual RNAseq: Human hepatocytes and Plasmodium berghei	Homo sapiens	Dual RNAseq - Human hepatocytes infected with Plasmodium berghei	PRJNA390648	52736152	99.44	Infected with Plasmodium berghei	NA
SRR5750526	Not	hepatocytes	SRX2950726		Dual RNAseq: Human hepatocytes and Plasmodium berghei	Homo sapiens	Dual RNAseq - Human hepatocytes infected with Plasmodium berghei	PRJNA390648	49216420	99.16	NA	NA
SRR6300667	Not	HUVEC	SRX3401581		GSM2859866: hyp_12h_n1_1; Homo sapiens; RNA-Seq	Homo sapiens	Total RNA deep sequencing (ribosomal depleted) of human umbilical vein endothelial cells exposed to hypoxia (0.2%) for 12h and 24h or kept under normoxic conditions.	PRJNA418883	39712730	94	Hypoxia 0.2% O2, 5% CO2 humidified atmosphere	
SRR6300668	Not	HUVEC	SRX3401582		GSM2859867: hyp_12h_n2_1; Homo sapiens; RNA-Seq	Homo sapiens	Total RNA deep sequencing (ribosomal depleted) of human umbilical vein endothelial cells exposed to hypoxia (0.2%) for 12h and 24h or kept under normoxic conditions.	PRJNA418883	43357452	94	Hypoxia 0.2% O2, 5% CO2 humidified atmosphere	
SRR6300669	Not	HUVEC	SRX3401583		GSM2859868: hyp_24h_n1_1; Homo sapiens; RNA-Seq	Homo sapiens	Total RNA deep sequencing (ribosomal depleted) of human umbilical vein endothelial cells exposed to hypoxia (0.2%) for 12h and 24h or kept under normoxic conditions.	PRJNA418883	41045511	94	Hypoxia 0.2% O2, 5% CO2 humidified atmosphere	
SRR6300670	Not	HUVEC	SRX3401584		GSM2859869: hyp_24h_n2_1; Homo sapiens; RNA-Seq	Homo sapiens	Total RNA deep sequencing (ribosomal depleted) of human umbilical vein endothelial cells exposed to hypoxia (0.2%) for 12h and 24h or kept under normoxic conditions.	PRJNA418883	44557122	94	Hypoxia 0.2% O2, 5% CO2 humidified atmosphere	
SRR6300671	Not	HUVEC	SRX3401585		GSM2859870: norm_12h_n1_1; Homo sapiens; RNA-Seq	Homo sapiens	Total RNA deep sequencing (ribosomal depleted) of human umbilical vein endothelial cells exposed to hypoxia (0.2%) for 12h and 24h or kept under normoxic conditions.	PRJNA418883	42304183	94	Normoxia 20% O2, 5% CO2 humidified atmosphere	
SRR6300672	Not	HUVEC	SRX3401586		GSM2859871: norm_12h_n2_1; Homo sapiens; RNA-Seq	Homo sapiens	Total RNA deep sequencing (ribosomal depleted) of human umbilical vein endothelial cells exposed to hypoxia (0.2%) for 12h and 24h or kept under normoxic conditions.	PRJNA418883	42748680	94	Normoxia 20% O2, 5% CO2 humidified atmosphere	
SRR6300673	Not	HUVEC	SRX3401587		GSM2859872: norm_24h_n1_1; Homo sapiens; RNA-Seq	Homo sapiens	Total RNA deep sequencing (ribosomal depleted) of human umbilical vein endothelial cells exposed to hypoxia (0.2%) for 12h and 24h or kept under normoxic conditions.	PRJNA418883	42771271	94	Normoxia 20% O2, 5% CO2 humidified atmosphere	
SRR6300674	Not	HUVEC	SRX3401588		GSM2859873: norm_24h_n2_1; Homo sapiens; RNA-Seq	Homo sapiens	Total RNA deep sequencing (ribosomal depleted) of human umbilical vein endothelial cells exposed to hypoxia (0.2%) for 12h and 24h or kept under normoxic conditions.	PRJNA418883	42131852	94	Normoxia 20% O2, 5% CO2 humidified atmosphere	
SRR3724617	Not	Lymphocytes	SRX1881712		GSM2218995: Healthy1-CD4Memory; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	29296358	100	healthy	
SRR3724618	Not	Lymphocytes	SRX1881713		GSM2218996: Healthy1-CD4Naive; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	27598581	100	healthy	
SRR3724619	Not	Lymphocytes	SRX1881714		GSM2218997: Healthy1-CD8Memory; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	26351498	100	healthy	
SRR3724620	Not	Lymphocytes	SRX1881715		GSM2218998: Healthy1-CD8Naive; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	29464342	100	healthy	
SRR3724621	Not	Lymphocytes	SRX1881716		GSM2218999: Healthy2-CD4Memory; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	48167581	100	healthy	
SRR3724622	Not	Lymphocytes	SRX1881717		GSM2219000: Healthy2-CD4Naive; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	42220340	100	healthy	
SRR3724623	Not	Lymphocytes	SRX1881718		GSM2219001: Healthy2-CD8Memory; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	38563382	100	healthy	
SRR3724624	Not	Lymphocytes	SRX1881719		GSM2219002: Healthy2-CD8Naive; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	45267949	100	healthy	
SRR3724625	Not	Lymphocytes	SRX1881720		GSM2219003: Healthy3-CD4Memory; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	45930063	100	healthy	
SRR3724626	Not	Lymphocytes	SRX1881721		GSM2219004: Healthy3-CD4Naive; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	40731526	100	healthy	
SRR3724627	Not	Lymphocytes	SRX1881722		GSM2219005: Healthy3-CD8Memory; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	47542258	100	healthy	
SRR3724628	Not	Lymphocytes	SRX1881723		GSM2219006: Healthy3-CD8Naive; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	53718245	100	healthy	
SRR3724629	Not	Lymphocytes	SRX1881724		GSM2219007: PNH1-CD4Memory; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	30025375	100	paroxysmal nocturnal hemoglobinuria (PNH)	
SRR3724630	Not	Lymphocytes	SRX1881725		GSM2219008: PNH1-CD4Naive; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	29638535	100	paroxysmal nocturnal hemoglobinuria (PNH)	
SRR3724631	Not	Lymphocytes	SRX1881726		GSM2219009: PNH1-CD8Memory; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	30553368	100	paroxysmal nocturnal hemoglobinuria (PNH)	
SRR3724632	Not	Lymphocytes	SRX1881727		GSM2219010: PNH1-CD8Naive; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	34722293	100	paroxysmal nocturnal hemoglobinuria (PNH)	

Table_S1 Metadata for differential analysis

SRR3724633	Not	Lymphocytes	SRX1881728	GSM2219011: PNH2-CD4Memory; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	32466930	100	paroxysmal nocturnal hemoglobinuria (PNH)
SRR3724634	Not	Lymphocytes	SRX1881729	GSM2219012: PNH2-CD4Naive; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	29139809	100	paroxysmal nocturnal hemoglobinuria (PNH)
SRR3724635	Not	Lymphocytes	SRX1881730	GSM2219013: PNH2-CD8Memory; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	28506728	100	paroxysmal nocturnal hemoglobinuria (PNH)
SRR3724636	Not	Lymphocytes	SRX1881731	GSM2219014: PNH2-CD8Naive; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	30666327	100	paroxysmal nocturnal hemoglobinuria (PNH)
SRR3724637	Not	Lymphocytes	SRX1881732	GSM2219015: PNH3-CD4Memory; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	49485440	100	paroxysmal nocturnal hemoglobinuria (PNH)
SRR3724638	Not	Lymphocytes	SRX1881733	GSM2219016: PNH3-CD4Naive; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	43428307	100	paroxysmal nocturnal hemoglobinuria (PNH)
SRR3724639	Not	Lymphocytes	SRX1881734	GSM2219017: PNH3-CD8Memory; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	41998683	100	paroxysmal nocturnal hemoglobinuria (PNH)
SRR3724640	Not	Lymphocytes	SRX1881735	GSM2219018: PNH3-CD8Naive; Homo sapiens; RNA-Seq	Homo sapiens	T cells from paroxysmal nocturnal hemoglobinuria patients show an altered TNFR signaling pathway	PRJNA327044	52622177	100	paroxysmal nocturnal hemoglobinuria (PNH)

Table_S2 Metadata for ENCODE kmer research

Original files	label	group	Accession	Assay Nickname	Biosample.summary	Lab	Project	Species	Life.stage	Age	Age.Units	Treatment
ENCF371KWE	dendritic cell	6: Others	ENCSR892ZUK	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 2 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF534HFO	dendritic cell	6: Others	ENCSR519NFO	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 1 hour	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF126HPJ	endothelial microvascular endothelial cells	8: Endothelial cells	ENCSR919MZM	total RNA-seq	endothelial microvascular endothelial cells female adult (34 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		34year	
ENCF502JCS	endothelial microvascular endothelial cells	8: Endothelial cells	ENCSR919MZM	total RNA-seq	endothelial microvascular endothelial cells female adult (34 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		34year	
ENCF001QZS	smooth muscle cell of bladder	7: SMC	ENCSR000AAC	total RNA-seq	smooth muscle cell of bladder female adult (53 years) and male adult (62 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		53.62year	
ENCF001RAO	smooth muscle cell of bladder	7: SMC	ENCSR000AAC	total RNA-seq	smooth muscle cell of bladder female adult (53 years) and male adult (62 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		53.62year	
ENCF001QZC	renal cortical epithelial cell	9: Epithelial cells	ENCSR000AAQ	total RNA-seq	renal cortical epithelial cell female adult (69 years) and male adult (84 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		69.84year	
ENCF001RAU	renal cortical epithelial cell	9: Epithelial cells	ENCSR000AAQ	total RNA-seq	renal cortical epithelial cell female adult (69 years) and male adult (84 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		69.84year	
ENCF001QZK	bladder microvascular endothelial cell	8: Endothelial cells	ENCSR000AAB	total RNA-seq	bladder microvascular endothelial cell male adult (46 years) and male adult (60 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		46.6year	
ENCF001RAI	bladder microvascular endothelial cell	8: Endothelial cells	ENCSR000AAB	total RNA-seq	bladder microvascular endothelial cell male adult (46 years) and male adult (60 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		46.6year	
ENCF001QZG	dermis blood vessel endothelial cell	8: Endothelial cells	ENCSR000AAI	total RNA-seq	dermis blood vessel endothelial cell female child (16 years) and male child (13 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	child		13.16year	
ENCF001RAQ	dermis blood vessel endothelial cell	8: Endothelial cells	ENCSR000AAI	total RNA-seq	dermis blood vessel endothelial cell female child (16 years) and male child (13 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	child		13.16year	
ENCF0000ILE	skeletal muscle satellite cell	4: Progenitor or Stem cell	ENCSR000CUI	total RNA-seq	skeletal muscle satellite cell female adult (64 years) and male adult (21 year)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		21.64year	
ENCF0000ILQ	skeletal muscle satellite cell	4: Progenitor or Stem cell	ENCSR000CUI	total RNA-seq	skeletal muscle satellite cell female adult (64 years) and male adult (21 year)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		21.64year	
ENCF0000GFS	mononuclear cell	6: Others	ENCSR000CUT	total RNA-seq	mononuclear cell female adult (52 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		52year	
ENCF810KXA	dendritic cell	6: Others	ENCSR682CFV	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 2 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF928LQX	dendritic cell	6: Others	ENCSR652RSO	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 6 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF816SSU	dendritic cell	6: Others	ENCSR536GUD	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 4 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF720QTK	dendritic cell	6: Others	ENCSR870DRE	total RNA-seq	dendritic cell treated with 0 ng/mL Lipopolysaccharide for 0 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF816QOY	H1-hESC	4: Progenitor or Stem cell	ENCSR537BCG	total RNA-seq	H1-hESC	Joe Ecker, Salk	Roadmap	Homo sapiens	embryonic	unknown	unknown	
ENCF476PHC	H1-hESC	4: Progenitor or Stem cell	ENCSR537BCG	total RNA-seq	H1-hESC	Joe Ecker, Salk	Roadmap	Homo sapiens	embryonic	unknown	unknown	
ENCF821YWS	H1-hESC	4: Progenitor or Stem cell	ENCSR537BCG	total RNA-seq	H1-hESC	Joe Ecker, Salk	Roadmap	Homo sapiens	embryonic	unknown	unknown	
ENCF712SHP	H1-hESC	4: Progenitor or Stem cell	ENCSR537BCG	total RNA-seq	H1-hESC	Joe Ecker, Salk	Roadmap	Homo sapiens	embryonic	unknown	unknown	
ENCF793PLC	H1-hESC	4: Progenitor or Stem cell	ENCSR537BCG	total RNA-seq	H1-hESC	Joe Ecker, Salk	Roadmap	Homo sapiens	embryonic	unknown	unknown	
ENCF000FKR	thoracic aorta endothelial cell	8: Endothelial cells	ENCSR000CUK	total RNA-seq	thoracic aorta endothelial cell female adult (22 years) and male adult (55 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		55.22year	
ENCF000FLI	thoracic aorta endothelial cell	8: Endothelial cells	ENCSR000CUK	total RNA-seq	thoracic aorta endothelial cell female adult (22 years) and male adult (55 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		55.22year	
ENCF000GBQ	hair follicle dermal papilla cell	4: Progenitor or Stem cell	ENCSR000CUB	total RNA-seq	hair follicle dermal papilla cell female adult (47 years) and female adult (70 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		70.47year	
ENCF000GBT	hair follicle dermal papilla cell	4: Progenitor or Stem cell	ENCSR000CUB	total RNA-seq	hair follicle dermal papilla cell female adult (47 years) and female adult (70 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		70.47year	
ENCF656QH	myometrial cell	5: Muscle cells	ENCSR371GVG	total RNA-seq	myometrial cell female adult (34 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		34year	
ENCF212TDB	myometrial cell	5: Muscle cells	ENCSR371GVG	total RNA-seq	myometrial cell female adult (34 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		34year	
ENCF995TMT	dendritic cell	6: Others	ENCSR332SRS	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 4 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF605VBT	glomerular endothelial cell	8: Endothelial cells	ENCSR878EUT	total RNA-seq	glomerular endothelial cell female embryo (22 weeks) and male embryo (22 weeks)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic		22week	
ENCF509JLH	glomerular endothelial cell	8: Endothelial cells	ENCSR878EUT	total RNA-seq	glomerular endothelial cell female embryo (22 weeks) and male embryo (22 weeks)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic		22week	
ENCF000FLO	articular chondrocyte	character(0)	ENCSR000CUE	total RNA-seq	articular chondrocyte of knee joint female adult (56 years) and male adult (64 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		64.56year	
ENCF000FMM	articular chondrocyte	character(0)	ENCSR000CUE	total RNA-seq	articular chondrocyte of knee joint female adult (56 years) and male adult (64 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		64.56year	
ENCF875NIU	dendritic cell	6: Others	ENCSR178SNP	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 1 hour	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF000GXC	fibroblast of villous mesenchyme	1: Fibroblast	ENCSR000CUL	total RNA-seq	fibroblast of villous mesenchyme female newborn and male newborn	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	newborn		unknown	
ENCF000GXV	fibroblast of villous mesenchyme	1: Fibroblast	ENCSR000CUL	total RNA-seq	fibroblast of villous mesenchyme female newborn and male newborn	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	newborn		unknown	
ENCF236EYN	hair follicular keratinocyte	6: Others	ENCSR680USE	total RNA-seq	hair follicular keratinocyte male adult (55 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		55year	
ENCF917MHA	hair follicular keratinocyte	6: Others	ENCSR680USE	total RNA-seq	hair follicular keratinocyte male adult (55 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		55year	
ENCF142YQX	mesangial cell	6: Others	ENCSR198TKA	total RNA-seq	mesangial cell NONE and female embryo (21 week)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown,embryo	unknown,21	week	
ENCF644UKX	mesangial cell	6: Others	ENCSR198TKA	total RNA-seq	mesangial cell NONE and female embryo (21 week)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown,embryo	unknown,21	week	
ENCF927MX	myotube	5: Muscle cells	ENCSR828TEI	total RNA-seq	myotube originated from skeletal muscle myoblast	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF999QZD	myotube	5: Muscle cells	ENCSR828TEI	total RNA-seq	myotube originated from skeletal muscle myoblast	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF911ELB	foreskin keratinocyte	6: Others	ENCSR034RPU	total RNA-seq	foreskin keratinocyte male newborn (2-4 days)	Michael Snyder, Stanford	GGR	Homo sapiens	newborn	2-4	day	
ENCF578BCTE	foreskin keratinocyte	6: Others	ENCSR034RPU	total RNA-seq	foreskin keratinocyte male newborn (2-4 days)	Michael Snyder, Stanford	GGR	Homo sapiens	newborn	2-4	day	
ENCF000GNB	vein endothelial cell	8: Endothelial cells	ENCSR000CUG	total RNA-seq	vein endothelial cell male adult (48 years) and male adult (52 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		52.48year	

Table_S2 Metadata for ENCODE kmer research

ENCF000GNC	vein endothelial cell	8: Endothelial cells	ENCSR000CUG	total RNA-seq	vein endothelial cell male adult (48 years) and male adult (52 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	52.48	year	
ENCF001QZM	smooth muscle cell of trachea	7: SMC	ENCSR000AAS	total RNA-seq	smooth muscle cell of trachea male adult (28 years) and male adult (56 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	56.28	year	
ENCF001RDQ	smooth muscle cell of trachea	7: SMC	ENCSR000AAS	total RNA-seq	smooth muscle cell of trachea male adult (28 years) and male adult (56 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	56.28	year	
ENCF001RCA	smooth muscle cell of the pulmonary artery	7: SMC	ENCSR000AAN	total RNA-seq	smooth muscle cell of the pulmonary artery male adult (26 years) and male adult (28 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	28.26	year	
ENCF001RDO	smooth muscle cell of the pulmonary artery	7: SMC	ENCSR000AAN	total RNA-seq	smooth muscle cell of the pulmonary artery male adult (26 years) and male adult (28 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	28.26	year	
ENCF000GKX	placental pericyte	6: Others	ENCSR000CTX	total RNA-seq	placental pericyte female newborn and male newborn	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	newborn	unknown		
ENCF000GKZ	placental pericyte	6: Others	ENCSR000CTX	total RNA-seq	placental pericyte female newborn and male newborn	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	newborn	unknown		
ENCF000HWI	fibroblast of dermis	1: Fibroblast	ENCSR000CUH	total RNA-seq	fibroblast of dermis female adult (44 years) and female adult (55 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	55.44	year	
ENCF000HXB	fibroblast of dermis	1: Fibroblast	ENCSR000CUH	total RNA-seq	fibroblast of dermis female adult (44 years) and female adult (55 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	55.44	year	
ENCF644CJJ	bipolar neuron	6: Others	ENCSR968WKR	total RNA-seq	bipolar neuron originated from GM23338 treated with 0.5 µg/mL doxycycline hyclate for 4 days	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	53	year	doxycycline hyclate
ENCF040TFC	bipolar neuron	6: Others	ENCSR968WKR	total RNA-seq	bipolar neuron originated from GM23338 treated with 0.5 µg/mL doxycycline hyclate for 4 days	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	53	year	doxycycline hyclate
ENCF000GFU	MSC of adipose	character(0)	ENCSR000CTZ	total RNA-seq	mesenchymal stem cell of adipose	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	42.37	year	
ENCF000GGR	MSC of adipose	character(0)	ENCSR000CTZ	total RNA-seq	mesenchymal stem cell of adipose	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	42.37	year	
ENCF001RBG	lung microvascular endothelial cell	8: Endothelial cells	ENCSR000AAP	total RNA-seq	lung microvascular endothelial cell female adult (55 years) and male adult (63 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	63.55	year	
ENCF001RBW	lung microvascular endothelial cell	8: Endothelial cells	ENCSR000AAP	total RNA-seq	lung microvascular endothelial cell female adult (55 years) and male adult (63 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	63.55	year	
ENCF840QXH	bronchus fibroblast of lung	1: Fibroblast	ENCSR620NSN	total RNA-seq	bronchus fibroblast of lung	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF480YLH	bronchus fibroblast of lung	1: Fibroblast	ENCSR620NSN	total RNA-seq	bronchus fibroblast of lung	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF001RAK	tracheal epithelial cell	9: Epithelial cells	ENCSR000AAR	total RNA-seq	tracheal epithelial cell male adult (21 year) and male adult (68 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	21.68	year	
ENCF001RBS	tracheal epithelial cell	9: Epithelial cells	ENCSR000AAR	total RNA-seq	tracheal epithelial cell male adult (21 year) and male adult (68 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	21.68	year	
ENCF001QZE	epithelial cell of viscerocranial mucosa	character(0)	ENCSR000AAL	total RNA-seq	nasal cavity respiratory epithelium epithelial cell of viscerocranial mucosa female adult (70 years) and male adult (46 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	46.7	year	
ENCF001QZI	epithelial cell of viscerocranial mucosa	character(0)	ENCSR000AAL	total RNA-seq	nasal cavity respiratory epithelium epithelial cell of viscerocranial mucosa female adult (70 years) and male adult (46 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	46.7	year	
ENCF001RBA	uterine smooth muscle cell	7: SMC	ENCSR000AAV	total RNA-seq	uterine smooth muscle cell female adult (48 years) and female adult (50 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	48.5	year	
ENCF001RCC	uterine smooth muscle cell	7: SMC	ENCSR000AAV	total RNA-seq	uterine smooth muscle cell female adult (48 years) and female adult (50 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	48.5	year	
ENCF001QZQ	aortic SMC	character(0)	ENCSR000AAA	total RNA-seq	aortic smooth muscle cell male adult (21 year) and male adult (54 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	54.21	year	
ENCF001RAM	aortic SMC	character(0)	ENCSR000AAA	total RNA-seq	aortic smooth muscle cell male adult (21 year) and male adult (54 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	54.21	year	
ENCF000IGA	melanocyte of skin	6: Others	ENCSR000CUQ	total RNA-seq	melanocyte of skin male child (1 year) and male child (3 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	child	3.1	year	
ENCF000IGU	melanocyte of skin	6: Others	ENCSR000CUQ	total RNA-seq	melanocyte of skin male child (1 year) and male child (3 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	child	3.1	year	
ENCF446LHV	dendritic cell	6: Others	ENCSR022BYE	total RNA-seq	dendritic cell treated with 0 ng/mL Lipopolysaccharide for 0 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF817RIH	dendritic cell	6: Others	ENCSR227MWL	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 2 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF662HMZ	dendritic cell	6: Others	ENCSR975YGW	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 2 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF001RAG	bronchial smooth muscle cell	7: SMC	ENCSR000AAE	total RNA-seq	bronchial smooth muscle cell male adult (52 years) and male adult (59 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	52.59	year	
ENCF001RAS	bronchial smooth muscle cell	7: SMC	ENCSR000AAE	total RNA-seq	bronchial smooth muscle cell male adult (52 years) and male adult (59 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	52.59	year	
ENCF001RAA	endothelial cell of coronary artery	8: Endothelial cells	ENCSR000AAF	total RNA-seq	endothelial cell of coronary artery female adult (41 year) and male adult (77 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	41.77	year	
ENCF001RBK	endothelial cell of coronary artery	8: Endothelial cells	ENCSR000AAF	total RNA-seq	endothelial cell of coronary artery female adult (41 year) and male adult (77 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	41.77	year	
ENCF001QZY	smooth muscle cell of the coronary artery	7: SMC	ENCSR000AAG	total RNA-seq	smooth muscle cell of the coronary artery female adult (53 years) and male adult (55 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	53.55	year	
ENCF001RCG	smooth muscle cell of the coronary artery	7: SMC	ENCSR000AAG	total RNA-seq	smooth muscle cell of the coronary artery female adult (53 years) and male adult (55 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	53.55	year	
ENCF001QZU	pulmonary artery endothelial cell	8: Endothelial cells	ENCSR000AAM	total RNA-seq	pulmonary artery endothelial cell male adult (23 years) and male adult (52 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	23.52	year	
ENCF001RCE	pulmonary artery endothelial cell	8: Endothelial cells	ENCSR000AAM	total RNA-seq	pulmonary artery endothelial cell male adult (23 years) and male adult (52 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	23.52	year	
ENCF071TOB	dendritic cell	6: Others	ENCSR084JIA	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 1 hour	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF001RAE	regular cardiac myocyte	5: Muscle cells	ENCSR000AAH	total RNA-seq	regular cardiac myocyte female adult (51 year) and male adult (48 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	48.51	year	
ENCF001RBM	regular cardiac myocyte	5: Muscle cells	ENCSR000AAH	total RNA-seq	regular cardiac myocyte female adult (51 year) and male adult (48 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	48.51	year	
ENCF001RAY	dermis lymphatic vessel endothelial cell	8: Endothelial cells	ENCSR000AAJ	total RNA-seq	dermis lymphatic vessel endothelial cell female adult (45 years) and male child (6 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	child,adult	6.45	year	

Table_S2 Metadata for ENCODE kmer research

ENCF001RBQ	dermis lymphatic vessel endothelial cell	8: Endothelial cells	ENCSR000AAJ	total RNA-seq	dermis lymphatic vessel endothelial cell female adult (45 years) and male child (6 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	child_adult	6,45	year	
ENCF001RAW	dermis microvascular lymphatic vessel endothelial cell	8: Endothelial cells	ENCSR000AAK	total RNA-seq	dermis microvascular lymphatic vessel endothelial cell female adult (38 years) and female adult (64 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	38,64	year	
ENCF001RBE	dermis microvascular lymphatic vessel endothelial cell	8: Endothelial cells	ENCSR000AAK	total RNA-seq	dermis microvascular lymphatic vessel endothelial cell female adult (38 years) and female adult (64 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	38,64	year	
ENCF365NAH	foreskin keratinocyte	6: Others	ENCSR527SSD	total RNA-seq	foreskin keratinocyte male newborn (2-4 days) treated with 1.2 mM calcium for 5.5 days	Michael Snyder, Stanford	GGR	Homo sapiens	newborn	2-4	day	calcium
ENCF668JVG	foreskin keratinocyte	6: Others	ENCSR527SSD	total RNA-seq	foreskin keratinocyte male newborn (2-4 days) treated with 1.2 mM calcium for 5.5 days	Michael Snyder, Stanford	GGR	Homo sapiens	newborn	2-4	day	calcium
ENCF649ASL	dendritic cell	6: Others	ENCSR571IUZ	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 4 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF123JND	dendritic cell	6: Others	ENCSR760INS	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 6 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF945OFS	dendritic cell	6: Others	ENCSR670MXT	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 2 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF124RBL	dendritic cell	6: Others	ENCSR476TKU	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 4 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF365UNT	dendritic cell	6: Others	ENCSR475TVN	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 6 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF658PGS	dendritic cell	6: Others	ENCSR707JVU	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 4 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF609HPU	dendritic cell	6: Others	ENCSR916YUR	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 6 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF000GLZ	placental epithelial cell	9: Epithelial cells	ENCSR000CUP	total RNA-seq	placental epithelial cell female newborn and male newborn	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	newborn	unknown		
ENCF000GMA	placental epithelial cell	9: Epithelial cells	ENCSR000CUP	total RNA-seq	placental epithelial cell female newborn and male newborn	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	newborn	unknown		
ENCF000GYE	subcutaneous preadipocyte	4: Progenitor or Stem cell	ENCSR000CUM	total RNA-seq	subcutaneous preadipocyte female adult (62 years) and male adult (65 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	65,62	year	
ENCF000GZD	subcutaneous preadipocyte	4: Progenitor or Stem cell	ENCSR000CUM	total RNA-seq	subcutaneous preadipocyte female adult (62 years) and male adult (65 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	65,62	year	
ENCF947BAG	dendritic cell	6: Others	ENCSR146PLL	total RNA-seq	dendritic cell treated with 100 ng/mL Lipopolysaccharide for 4 hours	Manuel Garber, UMass	GGR	Homo sapiens	unknown	unknown		Lipopolysaccharide
ENCF659EDN	mammary microvascular endothelial cell	8: Endothelial cells	ENCSR815UVL	total RNA-seq	mammary microvascular endothelial cell female adult (26 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	26	year	
ENCF143UMB	mammary microvascular endothelial cell	8: Endothelial cells	ENCSR815UVL	total RNA-seq	mammary microvascular endothelial cell female adult (26 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	26	year	
ENCF001QZW	epithelial cell of umbilical artery	9: Epithelial cells	ENCSR000AAT	total RNA-seq	epithelial cell of umbilical artery female newborn and male newborn	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	newborn	unknown		
ENCF001RBC	epithelial cell of umbilical artery	9: Epithelial cells	ENCSR000AAT	total RNA-seq	epithelial cell of umbilical artery female newborn and male newborn	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	newborn	unknown		
ENCF001RBU	smooth muscle cell of the umbilical artery	7: SMC	ENCSR000AAU	total RNA-seq	smooth muscle cell of the umbilical artery female newborn and male newborn	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	newborn	unknown		
ENCF001RBY	smooth muscle cell of the umbilical artery	7: SMC	ENCSR000AAU	total RNA-seq	smooth muscle cell of the umbilical artery female newborn and male newborn	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	newborn	unknown		
ENCF002DMJ	H7-hESC	4: Progenitor or Stem cell	ENCSR490SQH	total RNA-seq	H7-hESC	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic	unknown		
ENCF002DMP	H7-hESC	4: Progenitor or Stem cell	ENCSR490SQH	total RNA-seq	H7-hESC	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic	unknown		
ENCF000IHC	melanocyte of skin	6: Others	ENCSR000CUR	total RNA-seq	melanocyte of skin female adult (52 years) and male adult (55 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	55,52	year	
ENCF000IHY	melanocyte of skin	6: Others	ENCSR000CUR	total RNA-seq	melanocyte of skin female adult (52 years) and male adult (55 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	55,52	year	
ENCF575SLU	cardiac ventricle fibroblast	1: Fibroblast	ENCSR369RVN	total RNA-seq	cardiac ventricle fibroblast NONE and male adult (18 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult,unknown	18,unknown	year	
ENCF157UXG	cardiac ventricle fibroblast	1: Fibroblast	ENCSR369RVN	total RNA-seq	cardiac ventricle fibroblast NONE and male adult (18 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult,unknown	18,unknown	year	
ENCF689YVG	pericardium fibroblast	1: Fibroblast	ENCSR362HMX	total RNA-seq	pericardium fibroblast NONE and female embryo (20 weeks)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic,unkn	20,unknown	week	
ENCF005EEW	pericardium fibroblast	1: Fibroblast	ENCSR362HMX	total RNA-seq	pericardium fibroblast NONE and female embryo (20 weeks)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic,unkn	20,unknown	week	
ENCF746KSI	epithelial cell of alveolus of lung	9: Epithelial cells	ENCSR897KTO	total RNA-seq	epithelial cell of alveolus of lung NONE and female embryo (21 week)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown,embry	unknown,21	week	
ENCF408FWI	epithelial cell of alveolus of lung	9: Epithelial cells	ENCSR897KTO	total RNA-seq	epithelial cell of alveolus of lung NONE and female embryo (21 week)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown,embry	unknown,21	week	
ENCF109IUU	kidney epithelial cell	9: Epithelial cells	ENCSR373BDG	total RNA-seq	kidney epithelial cell male embryo (22 weeks) and male newborn	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic,newt	22,unknown	week	
ENCF999ZER	kidney epithelial cell	9: Epithelial cells	ENCSR373BDG	total RNA-seq	kidney epithelial cell male embryo (22 weeks) and male newborn	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic,newt	22,unknown	week	
ENCF353STY	cardiac atrium fibroblast	1: Fibroblast	ENCSR110BDY	total RNA-seq	cardiac atrium fibroblast male child (2 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	child	2	year	
ENCF753GPU	cardiac atrium fibroblast	1: Fibroblast	ENCSR110BDY	total RNA-seq	cardiac atrium fibroblast male child (2 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	child	2	year	
ENCF000GJC	osteoblast	3: Osteo/Chondro lineage	ENCSR000CUF	total RNA-seq	osteoblast female adult (56 years) and male adult (62 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	56,62	year	
ENCF000GKC	osteoblast	3: Osteo/Chondro lineage	ENCSR000CUF	total RNA-seq	osteoblast female adult (56 years) and male adult (62 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	56,62	year	
ENCF000GET	mammary epithelial cell	9: Epithelial cells	ENCSR000CUN	total RNA-seq	mammary epithelial cell female adult (23 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult	23	year	
ENCF353SPR	cardiac muscle cell	5: Muscle cells	ENCSR379YAE	total RNA-seq	cardiac muscle cell originated from RUES2	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic	unknown		
ENCF030WXU	cardiac muscle cell	5: Muscle cells	ENCSR379YAE	total RNA-seq	cardiac muscle cell originated from RUES2	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic	unknown		
ENCF472QXK	foreskin keratinocyte	6: Others	ENCSR959LTT	total RNA-seq	foreskin keratinocyte male newborn (2-4 days) treated with 1.2 mM calcium for 2.5 days	Michael Snyder, Stanford	GGR	Homo sapiens	newborn	2-4	day	calcium
ENCF218VQQ	foreskin keratinocyte	6: Others	ENCSR959LTT	total RNA-seq	foreskin keratinocyte male newborn (2-4 days) treated with 1.2 mM calcium for 2.5 days	Michael Snyder, Stanford	GGR	Homo sapiens	newborn	2-4	day	calcium

Table_S2 Metadata for ENCODE kmer research

ENCF000GIC	MSC of Wharton's jelly	character(0)	ENCSR000CUO	total RNA-seq	mesenchymal stem cell of Wharton's jelly	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	newborn	unknown		
ENCF000GIV	MSC of Wharton's jelly	character(0)	ENCSR000CUO	total RNA-seq	mesenchymal stem cell of Wharton's jelly	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	newborn	unknown		
ENCF548JWS	smooth muscle cell	7: SMC	ENCSR052FJA	total RNA-seq	smooth muscle cell originated from H9	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic		5day	
ENCF199QMY	smooth muscle cell	7: SMC	ENCSR052FJA	total RNA-seq	smooth muscle cell originated from H9	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic		5day	
ENCF001RAC	bronchial epithelial cell	9: Epithelial cells	ENCSR000AAD	total RNA-seq	bronchial epithelial cell female adult (40 years) and male adult (68 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		68.4year	
ENCF001RBO	bronchial epithelial cell	9: Epithelial cells	ENCSR000AAD	total RNA-seq	bronchial epithelial cell female adult (40 years) and male adult (68 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		68.4year	
ENCF653EZE	hepatocyte	6: Others	ENCSR908ZAS	total RNA-seq	hepatocyte originated from H9	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic		5day	
ENCF245VTB	hepatocyte	6: Others	ENCSR908ZAS	total RNA-seq	hepatocyte originated from H9	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic		5day	
ENCF996KMK	neural progenitor cell	4: Progenitor or Stem cell	ENCSR244ISQ	total RNA-seq	neural progenitor cell originated from H9	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic		5day	
ENCF939FVE	neural progenitor cell	4: Progenitor or Stem cell	ENCSR244ISQ	total RNA-seq	neural progenitor cell originated from H9	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	embryonic		5day	
ENCF002DMH	skeletal muscle myoblast	5: Muscle cells	ENCSR444WHQ	total RNA-seq	skeletal muscle myoblast	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF002DML	skeletal muscle myoblast	5: Muscle cells	ENCSR444WHQ	total RNA-seq	skeletal muscle myoblast	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF119TIN	myocyte	5: Muscle cells	ENCSR894WMQ	total RNA-seq	myocyte originated from LHCN-M2	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		41year	
ENCF700EUG	myocyte	5: Muscle cells	ENCSR894WMQ	total RNA-seq	myocyte originated from LHCN-M2	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		41year	
ENCF000GHA	MSC of the bone marrow	character(0)	ENCSR000CUD	total RNA-seq	mesenchymal stem cell of the bone marrow	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		57.6year	
ENCF000GHX	MSC of the bone marrow	character(0)	ENCSR000CUD	total RNA-seq	mesenchymal stem cell of the bone marrow	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		57.6year	
ENCF464PFE	airway epithelial cell	9: Epithelial cells	ENCSR822SUG	total RNA-seq	airway epithelial cell	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF846YCY	airway epithelial cell	9: Epithelial cells	ENCSR822SUG	total RNA-seq	airway epithelial cell	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF541NUM	epithelial cell of proximal tubule	9: Epithelial cells	ENCSR118TVR	total RNA-seq	epithelial cell of proximal tubule	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF479JTM	epithelial cell of proximal tubule	9: Epithelial cells	ENCSR118TVR	total RNA-seq	epithelial cell of proximal tubule	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF000FJX	fibroblast of the aortic adventitia	1: Fibroblast	ENCSR000CUJ	total RNA-seq	fibroblast of the aortic adventitia female adult (24 years) and male adult (47 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		47.24year	
ENCF000FKI	fibroblast of the aortic adventitia	1: Fibroblast	ENCSR000CUJ	total RNA-seq	fibroblast of the aortic adventitia female adult (24 years) and male adult (47 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		47.24year	
ENCF000EUS	hematopoietic multipotent progenitor cell	4: Progenitor or Stem cell	ENCSR000CUA	total RNA-seq	hematopoietic multipotent progenitor cell	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF393GQZ	astrocyte	6: Others	ENCSR233JIT	total RNA-seq	astrocyte	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF424DEO	astrocyte	6: Others	ENCSR233JIT	total RNA-seq	astrocyte	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	unknown	unknown		
ENCF001QZO	fibroblast of lung	1: Fibroblast	ENCSR000AAO	total RNA-seq	fibroblast of lung female adult (83 years) and male adult (23 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		83.23year	
ENCF001RBI	fibroblast of lung	1: Fibroblast	ENCSR000AAO	total RNA-seq	fibroblast of lung female adult (83 years) and male adult (23 years)	Thomas Gingeras, CSHL	ENCODE	Homo sapiens	adult		83.23year	

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

Run ID	titles	experiment_type	fraction	run_no	target_gene_symbol	specie	study_title	Study ID	nreads	read_length	library_selection	perturb_id	cell_type_alias
SRR7294023	single-cell RNA-seq of human adipose-derived mesenchymal stem cells	10X genomics sequencing	NA	NA	NA	Homo sapiens	single-cell transcriptomic sequencing of 24,370 adipose-derived mesenchymal stem cells	PRJNA472816	176545104	149.75	PolyA	NA	Ad-MSCs
SRR7253635	single-cell RNA-seq of human adipose-derived mesenchymal stem cells	10X genomics sequencing	NA	NA	NA	Homo sapiens	Single-cell transcriptomic sequencing of 24,370 adipose-derived mesenchymal stem cells	PRJNA472816	265719530	149.99	PolyA	NA	Ad-MSCs
SRR7363187	single-cell RNA-seq of human adipose-derived mesenchymal stem cells	10X genomics sequencing	NA	NA	NA	Homo sapiens	single-cell transcriptomic sequencing of 24,370 adipose-derived mesenchymal stem cells	PRJNA472816	352873537	149.98	PolyA	NA	Ad-MSCs
Run ID	titles	experiment_type	fraction	run_no	target_gene_symbol	specie	study_title	Study ID	nreads	read_length	library_selection	perturb_id	cell_type_alias
DRR177156	RDhi10602_ATCACG.experiment	Reference	cytoplasmic	0	NA	Homo sapiens	FANTOM6 project	PRJDB7993	33503637	51	Inverse rRNA	NA	HDF (neonata)
DRR177155	RDhi10602_ACTTGA.experiment	Reference	cytoplasmic	0	NA	Homo sapiens	FANTOM6 project	PRJDB7993	30096528	51	Inverse rRNA	NA	HDF (neonata)
DRR177154	RDhi10601_TTAGCG.experiment	Reference	nuclear	0	NA	Homo sapiens	FANTOM6 project	PRJDB7993	35912547	51	Inverse rRNA	NA	HDF (neonata)
DRR177153	RDhi10601_GATCAG.experiment	Reference	nuclear	0	NA	Homo sapiens	FANTOM6 project	PRJDB7993	35267146	51	Inverse rRNA	NA	HDF (neonata)
DRR177152	RDhi10600_ATCACG.experiment	Reference	chromatin	0	NA	Homo sapiens	FANTOM6 project	PRJDB7993	33268855	51	Inverse rRNA	NA	HDF (neonata)
DRR177151	RDhi10600_ACTTGA.experiment	Reference	chromatin	0	NA	Homo sapiens	FANTOM6 project	PRJDB7993	34620706	51	Inverse rRNA	NA	HDF (neonata)
DRR176801	CNhi10765_ACC.experiment	Targeted	NA	14	RP11-54A9.1	Homo sapiens	FANTOM6 project	PRJDB7993	17143511	47	CAGE	ASO_G0257219_AD_05	HDF (neonata)
DRR176800	CNhi10764_TAC.experiment	Targeted	NA	13	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	22288472	47	CAGE	ASO_G0270141_AD_10	HDF (neonata)
DRR176799	CNhi10764_GCT.experiment	Negative control	NA	14	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16779827	47	CAGE	ASO_Lipo	HDF (neonata)
DRR176798	CNhi10764_GCG.experiment	Targeted	NA	14	RP5-886K2.3	Homo sapiens	FANTOM6 project	PRJDB7993	13412787	47	CAGE	ASO_G0236810_AD_02	HDF (neonata)
DRR176797	CNhi10764_CAC.experiment	Targeted	NA	14	RP5-886K2.3	Homo sapiens	FANTOM6 project	PRJDB7993	13768161	47	CAGE	ASO_G0236810_AD_02	HDF (neonata)
DRR176796	CNhi10764_ATG.experiment	Experiment control	NA	13	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	18335819	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR176731	CNhi10640_ACG.experiment	Targeted	NA	12	DNM3OS	Homo sapiens	FANTOM6 project	PRJDB7993	14547994	47	CAGE	ASO_G0230630_AD_04	HDF (neonata)
DRR176730	CNhi10640_ACC.experiment	Targeted	NA	12	EMX2OS	Homo sapiens	FANTOM6 project	PRJDB7993	16086917	47	CAGE	ASO_G0229847_AD_04	HDF (neonata)
DRR176729	CNhi10639_TAC.experiment	Experiment control	NA	12	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	14896166	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR176728	CNhi10639_GCT.experiment	Targeted	NA	12	TUG1	Homo sapiens	FANTOM6 project	PRJDB7993	14574794	47	CAGE	ASO_G0253352_AD_08	HDF (neonata)
DRR176727	CNhi10639_GCG.experiment	Targeted	NA	12	C11orf95	Homo sapiens	FANTOM6 project	PRJDB7993	12726968	47	CAGE	ASO_G0188070_AD_05	HDF (neonata)
DRR176726	CNhi10639_CAC.experiment	Targeted	NA	12	LINC00339	Homo sapiens	FANTOM6 project	PRJDB7993	13664802	47	CAGE	ASO_G0218510_AD_03	HDF (neonata)
DRR176725	CNhi10639_ATG.experiment	Targeted	NA	12	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	12597402	47	CAGE	ASO_G0270141_AD_10	HDF (neonata)
DRR176724	CNhi10639_AGT.experiment	Negative control	NA	12	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11157330	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176659	CNhi10630_TAC.experiment	Negative control	NA	12	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14568347	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176658	CNhi10630_GCT.experiment	Targeted	NA	12	LINC00339	Homo sapiens	FANTOM6 project	PRJDB7993	14061748	47	CAGE	ASO_G0218510_AD_06	HDF (neonata)
DRR176657	CNhi10630_GCG.experiment	Targeted	NA	12	RP11-417E7.1	Homo sapiens	FANTOM6 project	PRJDB7993	12365042	47	CAGE	ASO_G0223485_AD_06	HDF (neonata)
DRR176656	CNhi10630_CAC.experiment	Targeted	NA	12	C11orf95	Homo sapiens	FANTOM6 project	PRJDB7993	13926089	47	CAGE	ASO_G0188070_AD_05	HDF (neonata)
DRR176655	CNhi10630_ATG.experiment	Targeted	NA	12	ERVX13.1	Homo sapiens	FANTOM6 project	PRJDB7993	15897264	47	CAGE	ASO_G0260565_AD_06	HDF (neonata)
DRR176654	CNhi10630_AGT.experiment	Targeted	NA	12	RP11-417E7.1	Homo sapiens	FANTOM6 project	PRJDB7993	11413198	47	CAGE	ASO_G0223485_AD_04	HDF (neonata)
DRR176653	CNhi10630_ACG.experiment	Targeted	NA	12	RP6-99M1.2	Homo sapiens	FANTOM6 project	PRJDB7993	14016329	47	CAGE	ASO_G0270069_AD_07	HDF (neonata)
DRR176652	CNhi10630_ACC.experiment	Targeted	NA	12	RP11-834C11.4	Homo sapiens	FANTOM6 project	PRJDB7993	20414932	47	CAGE	ASO_G0250742_AD_04	HDF (neonata)
DRR176587	CNhi10605_GCT.experiment	Targeted	NA	10	NR2F1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	535389	47	CAGE	ASO_G0237187_01	HDF (neonata)
DRR176586	CNhi10605_GCG.experiment	Targeted	NA	10	RP6-109B7.3	Homo sapiens	FANTOM6 project	PRJDB7993	10382541	47	CAGE	ASO_G0241990_07	HDF (neonata)
DRR176585	CNhi10605_CAC.experiment	Targeted	NA	10	LINC00674	Homo sapiens	FANTOM6 project	PRJDB7993	13123013	47	CAGE	ASO_G0237854_05	HDF (neonata)
DRR176584	CNhi10605_ATG.experiment	Experiment control	NA	10	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	14008394	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR176583	CNhi10605_AGT.experiment	Targeted	NA	10	CKMT2-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	16349504	47	CAGE	ASO_G0247572_09	HDF (neonata)
DRR176582	CNhi10605_ACG.experiment	Targeted	NA	11	LINC00883	Homo sapiens	FANTOM6 project	PRJDB7993	14714634	47	CAGE	ASO_G0243701_01	HDF (neonata)
DRR176581	CNhi10605_ACC.experiment	Targeted	NA	11	TP53TG1	Homo sapiens	FANTOM6 project	PRJDB7993	16141896	47	CAGE	ASO_G0182165_AD_10	HDF (neonata)
DRR176580	CNhi10604_TAC.experiment	Targeted	NA	10	LINC00707	Homo sapiens	FANTOM6 project	PRJDB7993	12703583	47	CAGE	ASO_G0238266_03	HDF (neonata)
DRR176515	CNhi10584_ATG.experiment	Targeted	NA	9	RP11-95P2.1	Homo sapiens	FANTOM6 project	PRJDB7993	11169949	47	CAGE	ASO_G0262468_07	HDF (neonata)
DRR176514	CNhi10584_AGT.experiment	Targeted	NA	9	RP11-95P2.1	Homo sapiens	FANTOM6 project	PRJDB7993	13068899	47	CAGE	ASO_G0262468_04	HDF (neonata)
DRR176513	CNhi10584_ACG.experiment	Targeted	NA	9	LINC00667	Homo sapiens	FANTOM6 project	PRJDB7993	11886738	47	CAGE	ASO_G0263753_01	HDF (neonata)
DRR176512	CNhi10584_ACC.experiment	Targeted	NA	9	RP11-95P2.1	Homo sapiens	FANTOM6 project	PRJDB7993	19111059	47	CAGE	ASO_G0262468_02	HDF (neonata)
DRR176511	CNhi10583_TAC.experiment	Targeted	NA	9	RP11-333E1.1	Homo sapiens	FANTOM6 project	PRJDB7993	11345714	47	CAGE	ASO_G0261879_05	HDF (neonata)
DRR176510	CNhi10583_GCT.experiment	Targeted	NA	9	RP11-95P2.1	Homo sapiens	FANTOM6 project	PRJDB7993	10783973	47	CAGE	ASO_G0262468_02	HDF (neonata)
DRR176509	CNhi10583_GCG.experiment	Targeted	NA	9	RP11-553K8.5	Homo sapiens	FANTOM6 project	PRJDB7993	12579190	47	CAGE	ASO_G0261573_03	HDF (neonata)
DRR176508	CNhi10583_CAC.experiment	Targeted	NA	9	RP11-553K8.5	Homo sapiens	FANTOM6 project	PRJDB7993	13226029	47	CAGE	ASO_G0261573_02	HDF (neonata)
DRR176443	CNhi10543_AGT.experiment	Targeted	NA	5	CATG00000016989.1	Homo sapiens	FANTOM6 project	PRJDB7993	13134737	47	CAGE	ASO_C0020443_06	HDF (neonata)
DRR176442	CNhi10543_ACG.experiment	Targeted	NA	5	CATG00000079799.1	Homo sapiens	FANTOM6 project	PRJDB7993	11721046	47	CAGE	ASO_C0008586_03	HDF (neonata)
DRR176441	CNhi10543_ACC.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12082179	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176440	CNhi10542_TAC.experiment	Reference	NA	0	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11753926	47	CAGE	NA	HDF (neonata)
DRR176439	CNhi10542_GCT.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15072508	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176438	CNhi10542_CAC.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12572386	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176437	CNhi10542_ATG.experiment	Targeted	NA	6	CTC-338M12.4	Homo sapiens	FANTOM6 project	PRJDB7993	10514355	47	CAGE	ASO_G0233937_05	HDF (neonata)
DRR176436	CNhi10542_AGT.experiment	Targeted	NA	6	RP11-220I.1	Homo sapiens	FANTOM6 project	PRJDB7993	11847706	47	CAGE	ASO_G0233137_04	HDF (neonata)
DRR176371	CNhi10534_ACC.experiment	Targeted	NA	6	AC092295.7	Homo sapiens	FANTOM6 project	PRJDB7993	11750669	47	CAGE	ASO_G0233527_07	HDF (neonata)
DRR176370	CNhi10533_TAC.experiment	Targeted	NA	5	KTNI1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10753751	47	CAGE	ASO_G0186615_03	HDF (neonata)
DRR176369	CNhi10533_GCT.experiment	Targeted	NA	6	AC092295.7	Homo sapiens	FANTOM6 project	PRJDB7993	10343841	47	CAGE	ASO_G0233527_02	HDF (neonata)
DRR176368	CNhi10533_GCG.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	9051548	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176367	CNhi10533_CAC.experiment	Targeted	NA	5	RP11-296O14.3	Homo sapiens	FANTOM6 project	PRJDB7993	10389408	47	CAGE	ASO_G0203739_10	HDF (neonata)
DRR176366	CNhi10533_ATG.experiment	Targeted	NA	6	AC092295.7	Homo sapiens	FANTOM6 project	PRJDB7993	10291057	47	CAGE	ASO_G0233527_07	HDF (neonata)
DRR176365	CNhi10533_AGT.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	9027563	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176364	CNhi10533_CAC.experiment	Targeted	NA	6	US2111.14	Homo sapiens	FANTOM6 project	PRJDB7993	11119652	47	CAGE	ASO_G0232725_05	HDF (neonata)
DRR176299	CNhi10524_TAC.experiment	Targeted	NA	5	CATG00000020700.1	Homo sapiens	FANTOM6 project	PRJDB7993	19711422	47	CAGE	ASO_C0020984_09	HDF (neonata)
DRR176298	CNhi10524_GCT.experiment	Targeted	NA	4	CD99P1	Homo sapiens	FANTOM6 project	PRJDB7993	16834564	47	CAGE	ASO_G0223773_05	HDF (neonata)
DRR176297	CNhi10524_GCG.experiment	Targeted	NA	4	RP11-395B7.4	Homo sapiens	FANTOM6 project	PRJDB7993	12730668	47	CAGE	ASO_G0227053_03	HDF (neonata)
DRR176296	CNhi10524_CAC.experiment	Negative control	NA	4	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16433061	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176295	CNhi10524_ATG.experiment	Targeted	NA	4	CD99P1	Homo sapiens	FANTOM6 project	PRJDB7993	17651165	47	CAGE	ASO_G0223773_06	HDF (neonata)

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176294	CNhi10524	ACG.experiment	Targeted	NA	4	MF12-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	15835443	47	CAGE	ASO G0228109_06	HDF (neonata)
DRR176293	CNhi10524	ACC.experiment	Targeted	NA	5	CATG00000095991.1	Homo sapiens	FANTOM6 project	PRJDB7993	21374168	47	CAGE	ASO C0012015_07	HDF (neonata)
DRR176292	CNhi10523	TAC.experiment	Targeted	NA	4	LL22NC03-86G7.1	Homo sapiens	FANTOM6 project	PRJDB7993	19989870	47	CAGE	ASO G0224086_03	HDF (neonata)
DRR176227	CNhi10492	GCG.experiment	Targeted	NA	3	AC108488.3	Homo sapiens	FANTOM6 project	PRJDB7993	11544811	47	CAGE	ASO G0234171_01	HDF (neonata)
DRR176226	CNhi10492	CAC.experiment	Targeted	NA	3	LINC00886	Homo sapiens	FANTOM6 project	PRJDB7993	12142331	47	CAGE	ASO G0240875_10	HDF (neonata)
DRR176225	CNhi10492	ATG.experiment	Targeted	NA	3	AC007246.3	Homo sapiens	FANTOM6 project	PRJDB7993	16768824	47	CAGE	ASO G0231312_04	HDF (neonata)
DRR176224	CNhi10492	AGT.experiment	Targeted	NA	3	ZNF131	Homo sapiens	FANTOM6 project	PRJDB7993	15842962	47	CAGE	ASO C0008543_05	HDF (neonata)
DRR176223	CNhi10492	ACG.experiment	Targeted	NA	3	AC007246.3	Homo sapiens	FANTOM6 project	PRJDB7993	14941536	47	CAGE	ASO G0231312_04	HDF (neonata)
DRR176222	CNhi10492	ACC.experiment	Targeted	NA	3	RP11-38L15.3	Homo sapiens	FANTOM6 project	PRJDB7993	14492220	47	CAGE	ASO G0231187_02	HDF (neonata)
DRR176221	CNhi10491	TAC.experiment	Targeted	NA	3	LINC00862	Homo sapiens	FANTOM6 project	PRJDB7993	17890185	47	CAGE	ASO G0203721_05	HDF (neonata)
DRR176220	CNhi10491	GCT.experiment	Targeted	NA	3	LINC00886	Homo sapiens	FANTOM6 project	PRJDB7993	16449157	47	CAGE	ASO G0240875_04	HDF (neonata)
DRR176155	CNhi10475	ATG.experiment	Targeted	NA	2	RP11-398K22.12	Homo sapiens	FANTOM6 project	PRJDB7993	19637512	47	CAGE	ASO G0229852_03	HDF (neonata)
DRR176154	CNhi10475	AGT.experiment	Targeted	NA	2	LINC00854	Homo sapiens	FANTOM6 project	PRJDB7993	16462388	47	CAGE	ASO G0205181_01	HDF (neonata)
DRR176153	CNhi10475	GCG.experiment	Targeted	NA	2	RP11-473M20.14	Homo sapiens	FANTOM6 project	PRJDB7993	19290770	47	CAGE	ASO G0263072_06	HDF (neonata)
DRR176152	CNhi10475	ACC.experiment	Targeted	NA	2	LINC00854	Homo sapiens	FANTOM6 project	PRJDB7993	18870242	47	CAGE	ASO G0205181_01	HDF (neonata)
DRR176151	CNhi10474	TAC.experiment	Targeted	NA	2	RP11-115C21.2	Homo sapiens	FANTOM6 project	PRJDB7993	18498643	47	CAGE	ASO G0246089_02	HDF (neonata)
DRR176150	CNhi10474	GCT.experiment	Negative control	NA	2	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16385749	47	CAGE	ASO NC A	HDF (neonata)
DRR176149	CNhi10474	GCG.experiment	Targeted	NA	2	NEAT1_2	Homo sapiens	FANTOM6 project	PRJDB7993	13959063	47	CAGE	ASO NEAT1_2	HDF (neonata)
DRR176148	CNhi10474	CAC.experiment	Targeted	NA	2	AC016747.3	Homo sapiens	FANTOM6 project	PRJDB7993	13658057	47	CAGE	ASO G0212978_04	HDF (neonata)
DRR176083	CNhi10466	AGT.experiment	Targeted	NA	2	AC005592.2	Homo sapiens	FANTOM6 project	PRJDB7993	16326963	47	CAGE	ASO G0231185_03	HDF (neonata)
DRR176082	CNhi10466	ACG.experiment	Targeted	NA	2	AC016747.3	Homo sapiens	FANTOM6 project	PRJDB7993	14489974	47	CAGE	ASO G0212978_01	HDF (neonata)
DRR176081	CNhi10466	ACC.experiment	Targeted	NA	2	AC016747.3	Homo sapiens	FANTOM6 project	PRJDB7993	22044906	47	CAGE	ASO G0212978_03	HDF (neonata)
DRR176080	CNhi10463	TAC.experiment	Targeted	NA	1	FGD5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	17735123	47	CAGE	ASO G0225733_04	HDF (neonata)
DRR176079	CNhi10463	GCT.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13264659	47	CAGE	ASO NC A	HDF (neonata)
DRR176078	CNhi10463	GCG.experiment	Targeted	NA	1	TUG1	Homo sapiens	FANTOM6 project	PRJDB7993	10978064	47	CAGE	ASO G0253352_06	HDF (neonata)
DRR176077	CNhi10463	CAC.experiment	Targeted	NA	1	LINC00938	Homo sapiens	FANTOM6 project	PRJDB7993	14501783	47	CAGE	ASO G0273015_05	HDF (neonata)
DRR176076	CNhi10463	ATG.experiment	Targeted	NA	1	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	15873018	47	CAGE	ASO G0226950_05	HDF (neonata)
DRR177150	CNhi10867	CAC.experiment	Targeted	NA	23	XRN2	Homo sapiens	FANTOM6 project	PRJDB7993	16021989	47	CAGE	siRNA siXRN-2	HDF (neonata)
DRR176180	CNhi10486	GCG.experiment	Targeted	NA	3	AC009948.5	Homo sapiens	FANTOM6 project	PRJDB7993	17124016	47	CAGE	ASO G0223960_08	HDF (neonata)
DRR176115	CNhi10470	ATG.experiment	Targeted	NA	2	LINC00630	Homo sapiens	FANTOM6 project	PRJDB7993	21992801	47	CAGE	ASO G0223546_01	HDF (neonata)
DRR176114	CNhi10470	AGT.experiment	Targeted	NA	2	RP11-611E13.2	Homo sapiens	FANTOM6 project	PRJDB7993	15920027	47	CAGE	ASO G0257815_01	HDF (neonata)
DRR176113	CNhi10470	ACC.experiment	Targeted	NA	2	BX322557.10	Homo sapiens	FANTOM6 project	PRJDB7993	15010004	47	CAGE	ASO G0215447_03	HDF (neonata)
DRR176112	CNhi10470	ACG.experiment	Targeted	NA	2	SERTAD4-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	23248785	47	CAGE	ASO G0203706_01	HDF (neonata)
DRR176111	CNhi10469	TAC.experiment	Targeted	NA	2	AC007879.7	Homo sapiens	FANTOM6 project	PRJDB7993	22846956	47	CAGE	ASO G0229647_06	HDF (neonata)
DRR176110	CNhi10469	GCT.experiment	Targeted	NA	2	RP11-458D21.1	Homo sapiens	FANTOM6 project	PRJDB7993	18718993	47	CAGE	ASO G0233396_02	HDF (neonata)
DRR176109	CNhi10469	GCG.experiment	Negative control	NA	2	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13142632	47	CAGE	ASO NC A	HDF (neonata)
DRR176108	CNhi10469	CAC.experiment	Targeted	NA	2	TRAM2-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	17294751	47	CAGE	ASO G0225791_05	HDF (neonata)
DRR177115	CNhi10836	ACG.experiment	Negative control	NA	18	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13655505	47	CAGE	ASO NC A	HDF (neonata)
DRR177114	CNhi10836	ACC.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	23739059	47	CAGE	ASO NC A	HDF (neonata)
DRR177113	CNhi10835	TAC.experiment	Targeted	NA	19	RP11-539L10.3	Homo sapiens	FANTOM6 project	PRJDB7993	25798059	47	CAGE	ASO G0251580_AD_05	HDF (neonata)
DRR177112	CNhi10835	GCT.experiment	Targeted	NA	18	CATG00000095991.1	Homo sapiens	FANTOM6 project	PRJDB7993	11567592	47	CAGE	ASO C0012015_AD_05	HDF (neonata)
DRR177111	CNhi10835	GCG.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13020861	47	CAGE	ASO NC B	HDF (neonata)
DRR177110	CNhi10835	CAC.experiment	Targeted	NA	18	CATG00000095991.1	Homo sapiens	FANTOM6 project	PRJDB7993	15861574	47	CAGE	ASO C0012015_AD_05	HDF (neonata)
DRR177109	CNhi10835	ATG.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14637155	47	CAGE	ASO NC B	HDF (neonata)
DRR177108	CNhi10835	AGT.experiment	Targeted	NA	21	JPX	Homo sapiens	FANTOM6 project	PRJDB7993	12147528	47	CAGE	ASO G0225470_AD_04	HDF (neonata)
DRR177043	CNhi10819	ACC.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15427139	47	CAGE	siRNA NC_2	HDF (neonata)
DRR177042	CNhi10818	TAC.experiment	Targeted	NA	20	LINC00630	Homo sapiens	FANTOM6 project	PRJDB7993	12929433	47	CAGE	siRNA n504056	HDF (neonata)
DRR177041	CNhi10818	GCT.experiment	Targeted	NA	21	LINC00959	Homo sapiens	FANTOM6 project	PRJDB7993	14818614	47	CAGE	ASO G0237489_AD_01	HDF (neonata)
DRR177040	CNhi10818	GCG.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14556123	47	CAGE	ASO Lipo	HDF (neonata)
DRR177039	CNhi10818	CAC.experiment	Targeted	NA	20	PRRX1	Homo sapiens	FANTOM6 project	PRJDB7993	10261513	47	CAGE	ASO G0116132_POS_09	HDF (neonata)
DRR177038	CNhi10818	ATG.experiment	Targeted	NA	20	CDKN2B-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	18746690	47	CAGE	ASO G0240498_POS_04	HDF (neonata)
DRR177037	CNhi10818	AGT.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	18459646	47	CAGE	ASO Untreat	HDF (neonata)
DRR177036	CNhi10818	ACG.experiment	Targeted	NA	21	HOTAIR	Homo sapiens	FANTOM6 project	PRJDB7993	17427449	47	CAGE	ASO G0228630_POS_06	HDF (neonata)
DRR176035	CNhi10458	ACC.experiment	Targeted	NA	1	AC017048.4	Homo sapiens	FANTOM6 project	PRJDB7993	18272029	47	CAGE	ASO G0224577_02	HDF (neonata)
DRR176034	CNhi10457	TAC.experiment	Targeted	NA	1	TUG1	Homo sapiens	FANTOM6 project	PRJDB7993	20936701	47	CAGE	ASO G0253352_06	HDF (neonata)
DRR176033	CNhi10457	GCT.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15933999	47	CAGE	ASO NC A	HDF (neonata)
DRR176032	CNhi10457	GCG.experiment	Targeted	NA	1	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	13039795	47	CAGE	ASO G0226950_05	HDF (neonata)
DRR176031	CNhi10457	CAC.experiment	Targeted	NA	1	AC017048.3	Homo sapiens	FANTOM6 project	PRJDB7993	13879212	47	CAGE	ASO G0163364_04	HDF (neonata)
DRR176030	CNhi10457	ATG.experiment	Targeted	NA	1	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	16115160	47	CAGE	ASO G0246067_05	HDF (neonata)
DRR176029	CNhi10457	AGT.experiment	Targeted	NA	1	LINC00702	Homo sapiens	FANTOM6 project	PRJDB7993	19006359	47	CAGE	ASO G0233117_04	HDF (neonata)
DRR176028	CNhi10457	ACG.experiment	Targeted	NA	1	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	17398562	47	CAGE	ASO G0226950_04	HDF (neonata)
DRR176971	CNhi10809	GCT.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15919384	47	CAGE	ASO Lipo	HDF (neonata)
DRR176970	CNhi10809	GCG.experiment	Targeted	NA	22	LINC00862	Homo sapiens	FANTOM6 project	PRJDB7993	11576336	47	CAGE	sgRNA blast5 LINC00862	HDF (neonata)
DRR176969	CNhi10809	CAC.experiment	Targeted	NA	22	LINC00630	Homo sapiens	FANTOM6 project	PRJDB7993	9961711	47	CAGE	sgRNA blast5 LINC00630	HDF (neonata)
DRR176968	CNhi10809	ATG.experiment	Targeted	NA	22	LINC00862	Homo sapiens	FANTOM6 project	PRJDB7993	17618203	47	CAGE	sgRNA blast5 LINC00862	HDF (neonata)
DRR176967	CNhi10809	AGT.experiment	Targeted	NA	22	LINC00630	Homo sapiens	FANTOM6 project	PRJDB7993	12091730	47	CAGE	sgRNA blast5 LINC00630	HDF (neonata)
DRR176966	CNhi10809	CAC.experiment	Experiment control	NA	20	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	24391514	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176965	CNhi10809	ACC.experiment	Targeted	NA	22	LINC00630	Homo sapiens	FANTOM6 project	PRJDB7993	18278971	47	CAGE	sgRNA blast5 LINC00630	HDF (neonata)
DRR176964	CNhi10808	TAC.experiment	Negative control	NA	22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	17737395	47	CAGE	sgRNA blast5 control	HDF (neonata)
DRR176899	CNhi10791	GCG.experiment	Targeted	NA	15	CTA-211A9.5	Homo sapiens	FANTOM6 project	PRJDB7993	15151803	47	CAGE	ASO G0244625_AD_09	HDF (neonata)
DRR176898	CNhi10791	CAC.experiment	Targeted	NA	15	CTD-2006C1.2	Homo sapiens	FANTOM6 project	PRJDB7993	14391634	47	CAGE	ASO G0219665_AD_04	HDF (neonata)
DRR176897	CNhi10791	AGT.experiment	Targeted	NA	16	RP11-6N17.4	Homo sapiens	FANTOM6 project	PRJDB7993	15366854	47	CAGE	ASO G0264920_AD_02	HDF (neonata)
DRR176896	CNhi10791	ACG.experiment	Experiment control	NA	16	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	20746443	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176895	CNhi10791	ACC.experiment	Targeted	NA	16	RP11-6N17.4	Homo sapiens	FANTOM6 project	PRJDB7993	22523908	47	CAGE	ASO G0264920_AD_02	HDF (neonata)
DRR176894	CNhi10790	TAC.experiment	Targeted	NA	16	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	22614587	47	CAGE	ASO G0226950_04	HDF (neonata)
DRR176893	CNhi10790	GCT.experiment	Experiment control	NA	15	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	15270427	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176892	CNhi10790	GCG.experiment	Targeted	NA	15	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	12986643	47	CAGE	ASO G0226950_04	HDF (neonata)
DRR176827	CNhi10782	ATG.experiment	Targeted	NA	16	LINC00667	Homo sapiens	FANTOM6 project	PRJDB7993	20781459	47	CAGE	ASO G0263753_AD_06	HDF (neonata)
DRR176826	CNhi10782	AGT.experiment	Negative control	NA	15	NA	Homo sapiens	FANTOM6 project	PRJDB7993	19734742	47	CAGE	ASO NC B	HDF (neonata)
DRR176825	CNhi10782	ACG.experiment	Targeted	NA	17	CTD-3018O17.3	Homo sapiens	FANTOM6 project	PRJDB7993	17599975	47	CAGE	ASO G0269834_AD_06	HDF (neonata)

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176824	CNhi10782	ACC.experiment	Negative control	NA	16	NA	Homo sapiens	FANTOM6 project	PRJDB7993	23277079	47	CAGE	ASO_Lipo		HDF (neonata)
DRR176823	CNhi10767	TAC.experiment	Negative control	NA	13	NA	Homo sapiens	FANTOM6 project	PRJDB7993	21972404	47	CAGE	ASO_Lipo		HDF (neonata)
DRR176822	CNhi10767	GCT.experiment	Targeted	NA	13	MEG3	Homo sapiens	FANTOM6 project	PRJDB7993	11757678	47	CAGE	ASO_G0214548	AD_06	HDF (neonata)
DRR176821	CNhi10767	GCG.experiment	Targeted	NA	14	MEG3	Homo sapiens	FANTOM6 project	PRJDB7993	10872425	47	CAGE	ASO_G0214548	AD_06	HDF (neonata)
DRR176820	CNhi10767	CAC.experiment	Targeted	NA	17	RP11-611E13.2	Homo sapiens	FANTOM6 project	PRJDB7993	16922148	47	CAGE	ASO_G0257815	AD_03	HDF (neonata)
DRR176755	CNhi10759	ACG.experiment	Targeted	NA	14	RP11-173B14.5	Homo sapiens	FANTOM6 project	PRJDB7993	19769317	47	CAGE	ASO_G0261105	AD_10	HDF (neonata)
DRR176754	CNhi10759	CAC.experiment	Targeted	NA	13	MAPKAPK5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	18228975	47	CAGE	ASO_G0234608	AD_01	HDF (neonata)
DRR176753	CNhi10758	TAC.experiment	Targeted	NA	13	RP11-527D7.1	Homo sapiens	FANTOM6 project	PRJDB7993	22627200	47	CAGE	ASO_G0225554	AD_04	HDF (neonata)
DRR176752	CNhi10758	GCT.experiment	Targeted	NA	13	RP11-139H15.1	Homo sapiens	FANTOM6 project	PRJDB7993	16023407	47	CAGE	ASO_G0225973	AD_01	HDF (neonata)
DRR176751	CNhi10758	GCG.experiment	Targeted	NA	14	LINC00517	Homo sapiens	FANTOM6 project	PRJDB7993	13517430	47	CAGE	ASO_G0259091	AD_04	HDF (neonata)
DRR176750	CNhi10758	CAC.experiment	Experiment control	NA	14	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	14146389	47	CAGE	ASO_MALAT1		HDF (neonata)
DRR176749	CNhi10758	ATG.experiment	Negative control	NA	13	NA	Homo sapiens	FANTOM6 project	PRJDB7993	8205374	47	CAGE	ASO_Untreat		HDF (neonata)
DRR176748	CNhi10758	AGT.experiment	Targeted	NA	13	MAPKAPK5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	18799556	47	CAGE	ASO_G0234608	AD_01	HDF (neonata)
DRR176683	CNhi10634	ACC.experiment	Targeted	NA	12	ERVK13-1	Homo sapiens	FANTOM6 project	PRJDB7993	16750405	47	CAGE	ASO_G0260565	AD_06	HDF (neonata)
DRR176682	CNhi10633	TAC.experiment	Targeted	NA	12	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	13894549	47	CAGE	ASO_G0230590	AD_03	HDF (neonata)
DRR176681	CNhi10633	GCT.experiment	Targeted	NA	12	ERVK3-1	Homo sapiens	FANTOM6 project	PRJDB7993	14931286	47	CAGE	ASO_G0124396	AD_01	HDF (neonata)
DRR176680	CNhi10633	GCG.experiment	Targeted	NA	12	LINC00339	Homo sapiens	FANTOM6 project	PRJDB7993	13046523	47	CAGE	ASO_G0218510	AD_06	HDF (neonata)
DRR176679	CNhi10633	CAC.experiment	Targeted	NA	12	MEG3	Homo sapiens	FANTOM6 project	PRJDB7993	15656972	47	CAGE	ASO_G0214548	AD_06	HDF (neonata)
DRR176678	CNhi10633	ATG.experiment	Negative control	NA	12	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13781324	47	CAGE	ASO_NC_A		HDF (neonata)
DRR176677	CNhi10633	AGT.experiment	Targeted	NA	12	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13191782	47	CAGE	ASO_G0246067	AD_09	HDF (neonata)
DRR176676	CNhi10633	ACG.experiment	Targeted	NA	12	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	14220102	47	CAGE	ASO_G0226950	AD_07	HDF (neonata)
DRR176611	CNhi10608	GCT.experiment	Targeted	NA	10	LINC00959	Homo sapiens	FANTOM6 project	PRJDB7993	13215357	47	CAGE	ASO_G0237489	09	HDF (neonata)
DRR176610	CNhi10608	GCG.experiment	Targeted	NA	10	ZSCAN16-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	7948507	47	CAGE	ASO_G0269293	01	HDF (neonata)
DRR176609	CNhi10608	CAC.experiment	Targeted	NA	10	ZNF730	Homo sapiens	FANTOM6 project	PRJDB7993	14389504	47	CAGE	ASO_G0267886	03	HDF (neonata)
DRR176608	CNhi10608	ATG.experiment	Targeted	NA	11	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	9486804	47	CAGE	ASO_G0246067	01	HDF (neonata)
DRR176607	CNhi10608	AGT.experiment	Targeted	NA	11	LINC00883	Homo sapiens	FANTOM6 project	PRJDB7993	11275756	47	CAGE	ASO_G0243701	AD_09	HDF (neonata)
DRR176606	CNhi10608	ACG.experiment	Targeted	NA	10	RP11-11011.12	Homo sapiens	FANTOM6 project	PRJDB7993	10956232	47	CAGE	ASO_G0255121	02	HDF (neonata)
DRR176605	CNhi10608	ACC.experiment	Experiment control	NA	10	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	10630460	47	CAGE	ASO_MALAT1		HDF (neonata)
DRR176604	CNhi10607	TAC.experiment	Experiment control	NA	10	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	12814570	47	CAGE	ASO_MALAT1		HDF (neonata)
DRR176539	CNhi10587	CAC.experiment	Targeted	NA	9	CTC-559E9.6	Homo sapiens	FANTOM6 project	PRJDB7993	11728992	47	CAGE	ASO_G0267419	02	HDF (neonata)
DRR176538	CNhi10587	ATG.experiment	Experiment control	NA	9	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	12599727	47	CAGE	ASO_MALAT1		HDF (neonata)
DRR176537	CNhi10587	AGT.experiment	Targeted	NA	9	AC009005.2	Homo sapiens	FANTOM6 project	PRJDB7993	8446828	47	CAGE	ASO_G0267751	07	HDF (neonata)
DRR176536	CNhi10587	ACG.experiment	Experiment control	NA	9	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	11244674	47	CAGE	ASO_MALAT1		HDF (neonata)
DRR176535	CNhi10587	ACC.experiment	Targeted	NA	9	CTC-559E9.6	Homo sapiens	FANTOM6 project	PRJDB7993	17399755	47	CAGE	ASO_G0267419	02	HDF (neonata)
DRR176534	CNhi10586	TAC.experiment	Targeted	NA	9	CTD-231912.4	Homo sapiens	FANTOM6 project	PRJDB7993	12397936	47	CAGE	ASO_G0267248	02	HDF (neonata)
DRR176533	CNhi10586	GCT.experiment	Targeted	NA	9	CTD-2162K18.5	Homo sapiens	FANTOM6 project	PRJDB7993	13818504	47	CAGE	ASO_G0267254	07	HDF (neonata)
DRR176532	CNhi10586	CAC.experiment	Targeted	NA	9	AC007773.2	Homo sapiens	FANTOM6 project	PRJDB7993	11299449	47	CAGE	ASO_G0267213	10	HDF (neonata)
DRR176467	CNhi10578	AGT.experiment	Targeted	NA	6	PITPNB-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	15144036	47	CAGE	ASO_G0236618	07	HDF (neonata)
DRR176466	CNhi10578	ACG.experiment	Experiment control	NA	6	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	12075213	47	CAGE	ASO_MALAT1		HDF (neonata)
DRR176465	CNhi10578	ACC.experiment	Targeted	NA	6	LINC00263	Homo sapiens	FANTOM6 project	PRJDB7993	14625012	47	CAGE	ASO_G0235823	03	HDF (neonata)
DRR176464	CNhi10577	TAC.experiment	Targeted	NA	6	RP11-357C3.3	Homo sapiens	FANTOM6 project	PRJDB7993	16395136	47	CAGE	ASO_G0235437	09	HDF (neonata)
DRR176463	CNhi10577	GCT.experiment	Targeted	NA	6	RP11-693N9.2	Homo sapiens	FANTOM6 project	PRJDB7993	13356982	47	CAGE	ASO_G0235505	01	HDF (neonata)
DRR176462	CNhi10577	GCG.experiment	Targeted	NA	6	RP11-357C3.3	Homo sapiens	FANTOM6 project	PRJDB7993	15790272	47	CAGE	ASO_G0235437	06	HDF (neonata)
DRR176461	CNhi10577	CAC.experiment	Targeted	NA	6	RP11-357C3.3	Homo sapiens	FANTOM6 project	PRJDB7993	15628154	47	CAGE	ASO_G0235437	05	HDF (neonata)
DRR176460	CNhi10577	ATG.experiment	Targeted	NA	6	RP11-357C3.3	Homo sapiens	FANTOM6 project	PRJDB7993	13198192	47	CAGE	ASO_G0235437	09	HDF (neonata)
DRR176395	CNhi10537	ACG.experiment	Targeted	NA	5	FLVCR1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13858935	47	CAGE	ASO_G0198468	03	HDF (neonata)
DRR176394	CNhi10537	ACC.experiment	Targeted	NA	6	CTC-338M12.4	Homo sapiens	FANTOM6 project	PRJDB7993	13596462	47	CAGE	ASO_G0233937	05	HDF (neonata)
DRR176393	CNhi10536	TAC.experiment	Targeted	NA	6	MAGI2-AS3	Homo sapiens	FANTOM6 project	PRJDB7993	14464102	47	CAGE	ASO_G0234456	04	HDF (neonata)
DRR176392	CNhi10536	GCT.experiment	Targeted	NA	5	FLVCR1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	17180713	47	CAGE	ASO_G0198468	01	HDF (neonata)
DRR176391	CNhi10536	GCG.experiment	Targeted	NA	5	CATG00000079799.1	Homo sapiens	FANTOM6 project	PRJDB7993	9728130	47	CAGE	ASO_C0008586	03	HDF (neonata)
DRR176390	CNhi10536	CAC.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11749411	47	CAGE	ASO_NC_A		HDF (neonata)
DRR176389	CNhi10536	AGT.experiment	Targeted	NA	5	FLVCR1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11365544	47	CAGE	ASO_G0198468	03	HDF (neonata)
DRR176388	CNhi10536	ACG.experiment	Targeted	NA	5	CATG00000016989.1	Homo sapiens	FANTOM6 project	PRJDB7993	11449770	47	CAGE	ASO_C0020443	03	HDF (neonata)
DRR176323	CNhi10527	TAC.experiment	Targeted	NA	4	SUGT1P	Homo sapiens	FANTOM6 project	PRJDB7993	19369060	47	CAGE	ASO_G0226823	08	HDF (neonata)
DRR176322	CNhi10527	GCT.experiment	Targeted	NA	5	CDKN2C	Homo sapiens	FANTOM6 project	PRJDB7993	17776122	47	CAGE	ASO_C0000830	02	HDF (neonata)
DRR176321	CNhi10527	GCG.experiment	Targeted	NA	4	RP11-395B7.4	Homo sapiens	FANTOM6 project	PRJDB7993	13714212	47	CAGE	ASO_G0227053	04	HDF (neonata)
DRR176320	CNhi10527	CAC.experiment	Targeted	NA	4	RP4-773A18.4	Homo sapiens	FANTOM6 project	PRJDB7993	18580280	47	CAGE	ASO_G0227811	06	HDF (neonata)
DRR176319	CNhi10527	ATG.experiment	Targeted	NA	5	CATG00000020700.1	Homo sapiens	FANTOM6 project	PRJDB7993	15653322	47	CAGE	ASO_C0020984	09	HDF (neonata)
DRR176318	CNhi10527	AGT.experiment	Targeted	NA	4	PCAT6	Homo sapiens	FANTOM6 project	PRJDB7993	18372099	47	CAGE	ASO_G0228288	01	HDF (neonata)
DRR176317	CNhi10527	ACG.experiment	Negative control	NA	4	NA	Homo sapiens	FANTOM6 project	PRJDB7993	10247717	47	CAGE	ASO_NC_A		HDF (neonata)
DRR176316	CNhi10527	ACC.experiment	Targeted	NA	5	CDKN2C	Homo sapiens	FANTOM6 project	PRJDB7993	18047930	47	CAGE	ASO_C0000830	08	HDF (neonata)
DRR176251	CNhi10495	GCT.experiment	Targeted	NA	3	LINC00886	Homo sapiens	FANTOM6 project	PRJDB7993	12730601	47	CAGE	ASO_G0204875	04	HDF (neonata)
DRR176250	CNhi10495	CAC.experiment	Targeted	NA	3	AC004980.7	Homo sapiens	FANTOM6 project	PRJDB7993	8280181	47	CAGE	ASO_G0205485	03	HDF (neonata)
DRR176249	CNhi10495	ATG.experiment	Targeted	NA	3	AC007246.3	Homo sapiens	FANTOM6 project	PRJDB7993	10763090	47	CAGE	ASO_G0231312	02	HDF (neonata)
DRR176248	CNhi10495	AGT.experiment	Targeted	NA	3	PRRT3-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11403597	47	CAGE	ASO_G0230082	01	HDF (neonata)
DRR176247	CNhi10495	ACG.experiment	Targeted	NA	3	CTD-2587H24.5	Homo sapiens	FANTOM6 project	PRJDB7993	10939951	47	CAGE	ASO_G0267577	03	HDF (neonata)
DRR176246	CNhi10495	ACC.experiment	Targeted	NA	3	LINC00856	Homo sapiens	FANTOM6 project	PRJDB7993	12788498	47	CAGE	ASO_G0230417	02	HDF (neonata)
DRR176245	CNhi10494	TAC.experiment	Targeted	NA	3	AC004980.7	Homo sapiens	FANTOM6 project	PRJDB7993	13771619	47	CAGE	ASO_G0205485	01	HDF (neonata)
DRR176244	CNhi10494	GCT.experiment	Negative control	NA	3	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12477515	47	CAGE	ASO_NC_A		HDF (neonata)
DRR176179	CNhi10486	CAC.experiment	Targeted	NA	3	CATG00000017883.1	Homo sapiens	FANTOM6 project	PRJDB7993	15328070	47	CAGE	ASO_C0020672	07	HDF (neonata)
DRR176178	CNhi10486	ATG.experiment	Targeted	NA	3	ZNF674-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14830624	47	CAGE	ASO_G0230844	01	HDF (neonata)
DRR176177	CNhi10486	AGT.experiment	Targeted	NA	3	FBXL19-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	15900573	47	CAGE	ASO_G0260852	02	HDF (neonata)
DRR176176	CNhi10486	ACG.experiment	Negative control	NA	3	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11850730	47	CAGE	ASO_NC_A		HDF (neonata)
DRR176175	CNhi10486	ACC.experiment	Targeted	NA	3	LINC00886	Homo sapiens	FANTOM6 project	PRJDB7993	17017772	47	CAGE	ASO_G0204875	01	HDF (neonata)
DRR176174	CNhi10477	TAC.experiment	Targeted	NA	2	RP11-150C12.1	Homo sapiens	FANTOM6 project	PRJDB7993	17106848	47	CAGE	ASO_C013368	01	HDF (neonata)
DRR176173	CNhi10477	GCT.experiment	Targeted	NA	2	CATG00000108369.1	Homo sapiens	FANTOM6 project	PRJDB7993	17651172	47	CAGE	ASO_C014559	06	HDF (neonata)
DRR176172	CNhi10477	GCG.experiment	Targeted	NA	2	RP11-65J3.1	Homo sapiens	FANTOM6 project	PRJDB7993	15182075	47	CAGE	ASO_G0233901	02	HDF (neonata)
DRR176107	CNhi10469	AGT.experiment	Targeted	NA	2	LINC00968	Homo sapiens	FANTOM6 project	PRJDB7993	13482789	47	CAGE	ASO_G0246430	03	HDF (neonata)

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176106	CNhi10469	ACG.experiment	Targeted	NA		2	CTD-3131K8.2	Homo sapiens	FANTOM6 project	PRJDB7993	17628008	47	CAGE	ASO G0269439_09	HDF (neonata)
DRR176105	CNhi10469	ACC.experiment	Targeted	NA		2	BX322557.10	Homo sapiens	FANTOM6 project	PRJDB7993	20027503	47	CAGE	ASO G0215447_01	HDF (neonata)
DRR176104	CNhi10468	TAC.experiment	Targeted	NA		2	CATG00000089639.1	Homo sapiens	FANTOM6 project	PRJDB7993	24217388	47	CAGE	ASO G0223811_06	HDF (neonata)
DRR176103	CNhi10468	GCT.experiment	Targeted	NA		2	RP11-150C12.1	Homo sapiens	FANTOM6 project	PRJDB7993	18033165	47	CAGE	ASO C013368_02	HDF (neonata)
DRR176102	CNhi10468	GCG.experiment	Experiment control	NA		2	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	13340172	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176101	CNhi10468	CAC.experiment	Targeted	NA		2	RP11-473M20.14	Homo sapiens	FANTOM6 project	PRJDB7993	15962914	47	CAGE	ASO G0263072_05	HDF (neonata)
DRR176100	CNhi10468	ATG.experiment	Targeted	NA		2	RP11-65J3.1	Homo sapiens	FANTOM6 project	PRJDB7993	18863550	47	CAGE	ASO G0233901_04	HDF (neonata)
DRR177107	CNhi10835	ACG.experiment	Targeted	NA		18	NA	Homo sapiens	FANTOM6 project	PRJDB7993	1729127	47	CAGE	ASO C0011632_AD_03	HDF (neonata)
DRR177106	CNhi10835	ACC.experiment	Negative control	NA		18	NA	Homo sapiens	FANTOM6 project	PRJDB7993	22721284	47	CAGE	ASO Unireat	HDF (neonata)
DRR177105	CNhi10834	TAC.experiment	Targeted	NA		19	NR2F1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	23083302	47	CAGE	ASO G0237187_AD_08	HDF (neonata)
DRR177104	CNhi10834	GCT.experiment	Targeted	NA		21	WDFY3-AS2	Homo sapiens	FANTOM6 project	PRJDB7993	9506419	47	CAGE	ASO G0180769_AD_03	HDF (neonata)
DRR177103	CNhi10834	GCG.experiment	Targeted	NA		21	TMCC1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10323399	47	CAGE	ASO G0271270_AD_10	HDF (neonata)
DRR177102	CNhi10834	CAC.experiment	Targeted	NA		19	CTC-228N24.3	Homo sapiens	FANTOM6 project	PRJDB7993	16179425	47	CAGE	ASO G0245937_AD_02	HDF (neonata)
DRR177101	CNhi10834	ATG.experiment	Targeted	NA		19	LINC00968	Homo sapiens	FANTOM6 project	PRJDB7993	22688373	47	CAGE	ASO G0246430_AD_03	HDF (neonata)
DRR177100	CNhi10834	AGT.experiment	Targeted	NA		18	CATG00000095991.1	Homo sapiens	FANTOM6 project	PRJDB7993	11682214	47	CAGE	ASO C0012015_AD_04	HDF (neonata)
DRR177035	CNhi10818	ACC.experiment	Targeted	NA		20	LINC00630	Homo sapiens	FANTOM6 project	PRJDB7993	4723849	47	CAGE	siRNA n504056	HDF (neonata)
DRR177034	CNhi10817	TAC.experiment	Targeted	NA		20	EMC7	Homo sapiens	FANTOM6 project	PRJDB7993	21757196	47	CAGE	siRNA s32293	HDF (neonata)
DRR177033	CNhi10817	GCT.experiment	Targeted	NA		20	PRRX1	Homo sapiens	FANTOM6 project	PRJDB7993	10102944	47	CAGE	siRNA s10745	HDF (neonata)
DRR177032	CNhi10817	GCG.experiment	Targeted	NA		17	FOXN3-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	12083637	47	CAGE	ASO G0258920_AD_07	HDF (neonata)
DRR177031	CNhi10817	CAC.experiment	Targeted	NA		20	EMC7	Homo sapiens	FANTOM6 project	PRJDB7993	19844065	47	CAGE	ASO G0134153_POS_05	HDF (neonata)
DRR177030	CNhi10817	ATG.experiment	Negative control	NA		21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14060457	47	CAGE	ASO NC_B	HDF (neonata)
DRR177029	CNhi10817	AGT.experiment	Targeted	NA		21	HOTAIR	Homo sapiens	FANTOM6 project	PRJDB7993	12355057	47	CAGE	ASO G0228630_POS_07	HDF (neonata)
DRR177028	CNhi10817	ACC.experiment	Negative control	NA		13	NA	Homo sapiens	FANTOM6 project	PRJDB7993	20891062	47	CAGE	ASO NC_B	HDF (neonata)
DRR176027	CNhi10457	ACC.experiment	Targeted	NA		1	AC017048.4	Homo sapiens	FANTOM6 project	PRJDB7993	20007183	47	CAGE	ASO G0224577_01	HDF (neonata)
DRR176026	CNhi10456	TAC.experiment	Negative control	NA		1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	23294421	47	CAGE	ASO NC_A	HDF (neonata)
DRR176025	CNhi10456	GCT.experiment	Targeted	NA		1	LINC00883	Homo sapiens	FANTOM6 project	PRJDB7993	37440	47	CAGE	ASO G0243701_01	HDF (neonata)
DRR176024	CNhi10456	GCG.experiment	Targeted	NA		1	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	17935999	47	CAGE	ASO G0230590_05	HDF (neonata)
DRR176023	CNhi10456	CAC.experiment	Targeted	NA		1	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	16500362	47	CAGE	ASO G0226950_02	HDF (neonata)
DRR176022	CNhi10456	ATG.experiment	Targeted	NA		1	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	22841354	47	CAGE	ASO G0230590_06	HDF (neonata)
DRR176021	CNhi10456	AGT.experiment	Targeted	NA		1	TUG1	Homo sapiens	FANTOM6 project	PRJDB7993	22896867	47	CAGE	ASO G0253352_02	HDF (neonata)
DRR176020	CNhi10456	ACG.experiment	Targeted	NA		1	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	19122181	47	CAGE	ASO G0270141_05	HDF (neonata)
DRR176963	CNhi10808	GCT.experiment	Negative control	NA		22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16040232	47	CAGE	sgRNA NC gRNA2	HDF (neonata)
DRR176962	CNhi10808	GCG.experiment	Negative control	NA		22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	10798557	47	CAGE	sgRNA NC	HDF (neonata)
DRR176961	CNhi10808	CAC.experiment	Targeted	NA		22	LINC00862	Homo sapiens	FANTOM6 project	PRJDB7993	10951182	47	CAGE	sgRNA blast5 LINC00862	HDF (neonata)
DRR176960	CNhi10808	ATG.experiment	Negative control	NA		22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	18657388	47	CAGE	sgRNA dCas9	HDF (neonata)
DRR176959	CNhi10808	AGT.experiment	Targeted	NA		22	LINC00862	Homo sapiens	FANTOM6 project	PRJDB7993	18296607	47	CAGE	sgRNA blast5 LINC00862	HDF (neonata)
DRR176958	CNhi10808	ACC.experiment	Negative control	NA		22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16340984	47	CAGE	sgRNA NC gRNA1	HDF (neonata)
DRR176957	CNhi10808	ACG.experiment	Targeted	NA		22	LINC00862	Homo sapiens	FANTOM6 project	PRJDB7993	14710450	47	CAGE	sgRNA blast5 LINC00862	HDF (neonata)
DRR176956	CNhi10807	TAC.experiment	Targeted	NA		22	LINC00630	Homo sapiens	FANTOM6 project	PRJDB7993	22091564	47	CAGE	sgRNA blast5 LINC00630	HDF (neonata)
DRR176891	CNhi10790	CAC.experiment	Targeted	NA		16	CTD-2319I12.4	Homo sapiens	FANTOM6 project	PRJDB7993	17435624	47	CAGE	ASO G0267248_AD_03	HDF (neonata)
DRR176890	CNhi10790	ATG.experiment	Targeted	NA		16	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	18138614	47	CAGE	ASO G0226950_04	HDF (neonata)
DRR176889	CNhi10790	AGT.experiment	Targeted	NA		16	CTD-2319I12.4	Homo sapiens	FANTOM6 project	PRJDB7993	11112769	47	CAGE	ASO G0267248_AD_03	HDF (neonata)
DRR176888	CNhi10790	ACG.experiment	Negative control	NA		15	NA	Homo sapiens	FANTOM6 project	PRJDB7993	18258451	47	CAGE	ASO NC_A	HDF (neonata)
DRR176887	CNhi10790	ACC.experiment	Targeted	NA		15	PITPNA-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	24549007	47	CAGE	ASO G0236618_AD_09	HDF (neonata)
DRR176886	CNhi10789	TAC.experiment	Negative control	NA		17	NA	Homo sapiens	FANTOM6 project	PRJDB7993	19085840	47	CAGE	ASO NC_B	HDF (neonata)
DRR176885	CNhi10789	GCT.experiment	Negative control	NA		15	NA	Homo sapiens	FANTOM6 project	PRJDB7993	17965221	47	CAGE	ASO NC_A	HDF (neonata)
DRR176884	CNhi10789	GCG.experiment	Targeted	NA		15	CTA-211A9.5	Homo sapiens	FANTOM6 project	PRJDB7993	15445531	47	CAGE	ASO G0244625_AD_09	HDF (neonata)
DRR176819	CNhi10767	AGT.experiment	Targeted	NA		14	MEG3	Homo sapiens	FANTOM6 project	PRJDB7993	12316424	47	CAGE	ASO G0214548_AD_06	HDF (neonata)
DRR176818	CNhi10767	ACG.experiment	Negative control	NA		17	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15284019	47	CAGE	ASO NC_A	HDF (neonata)
DRR176817	CNhi10767	CAC.experiment	Targeted	NA		14	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	17597100	47	CAGE	ASO G0226950_04	HDF (neonata)
DRR176816	CNhi10766	TAC.experiment	Targeted	NA		14	RAD51-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	20107651	47	CAGE	ASO G0245849_AD_07	HDF (neonata)
DRR176815	CNhi10766	GCT.experiment	Targeted	NA		14	LINC00517	Homo sapiens	FANTOM6 project	PRJDB7993	14788123	47	CAGE	ASO G0259091_AD_10	HDF (neonata)
DRR176814	CNhi10766	GCG.experiment	Negative control	NA		14	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14774541	47	CAGE	ASO NC_A	HDF (neonata)
DRR176813	CNhi10766	CAC.experiment	Negative control	NA		13	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14704305	47	CAGE	ASO NC_A	HDF (neonata)
DRR176812	CNhi10766	ATG.experiment	Negative control	NA		14	NA	Homo sapiens	FANTOM6 project	PRJDB7993	18819813	47	CAGE	ASO NC_A	HDF (neonata)
DRR176747	CNhi10758	ACG.experiment	Targeted	NA		13	RP11-295G20.2	Homo sapiens	FANTOM6 project	PRJDB7993	19693852	47	CAGE	ASO G0233461_AD_03	HDF (neonata)
DRR176746	CNhi10758	ACC.experiment	Targeted	NA		13	MAPKAPK5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	20424485	47	CAGE	ASO G0234608_AD_04	HDF (neonata)
DRR176745	CNhi10757	TAC.experiment	Targeted	NA		14	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	20703281	47	CAGE	ASO G0270141_AD_10	HDF (neonata)
DRR176744	CNhi10757	GCT.experiment	Targeted	NA		14	RP11-173B14.5	Homo sapiens	FANTOM6 project	PRJDB7993	14477585	47	CAGE	ASO G0261105_AD_06	HDF (neonata)
DRR176743	CNhi10757	GCG.experiment	Experiment control	NA		13	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	10319149	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176742	CNhi10757	CAC.experiment	Targeted	NA		13	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	14472240	47	CAGE	ASO G0226950_04	HDF (neonata)
DRR176741	CNhi10757	ATG.experiment	Targeted	NA		13	MAPKAPK5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	21528387	47	CAGE	ASO G0234608_AD_04	HDF (neonata)
DRR176740	CNhi10757	AGT.experiment	Targeted	NA		13	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	16450408	47	CAGE	ASO G0270141_AD_10	HDF (neonata)
DRR176675	CNhi10633	ACC.experiment	Targeted	NA		12	FGD5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	18575231	47	CAGE	ASO G0225733_AD_06	HDF (neonata)
DRR176674	CNhi10632	TAC.experiment	Targeted	NA		12	MEG3	Homo sapiens	FANTOM6 project	PRJDB7993	16270762	47	CAGE	ASO G0214548_AD_09	HDF (neonata)
DRR176673	CNhi10632	GCT.experiment	Targeted	NA		12	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	12692593	47	CAGE	ASO G0270141_AD_10	HDF (neonata)
DRR176672	CNhi10632	GCG.experiment	Targeted	NA		12	RP11-221N13.3	Homo sapiens	FANTOM6 project	PRJDB7993	11500230	47	CAGE	ASO G0256268_AD_07	HDF (neonata)
DRR176671	CNhi10632	CAC.experiment	Targeted	NA		12	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14255046	47	CAGE	ASO G0246067_AD_06	HDF (neonata)
DRR176670	CNhi10632	AGT.experiment	Targeted	NA		12	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	13725236	47	CAGE	ASO G0230590_AD_03	HDF (neonata)
DRR176669	CNhi10632	ACG.experiment	Targeted	NA		12	ERVX13-1	Homo sapiens	FANTOM6 project	PRJDB7993	13073268	47	CAGE	ASO G0260565_AD_08	HDF (neonata)
DRR176668	CNhi10632	ACC.experiment	Targeted	NA		12	MEG3	Homo sapiens	FANTOM6 project	PRJDB7993	16724551	47	CAGE	ASO G0214548_AD_06	HDF (neonata)
DRR176603	CNhi10607	GCT.experiment	Targeted	NA		10	RP13-463N16.6	Homo sapiens	FANTOM6 project	PRJDB7993	11747961	47	CAGE	ASO G0242147_07	HDF (neonata)
DRR176602	CNhi10607	GCG.experiment	Targeted	NA		10	NR2F1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10310715	47	CAGE	ASO G0237187_03	HDF (neonata)
DRR176601	CNhi10607	CAC.experiment	Targeted	NA		10	CTA-211A9.5	Homo sapiens	FANTOM6 project	PRJDB7993	12440397	47	CAGE	ASO G0244625_04	HDF (neonata)
DRR176600	CNhi10607	ATG.experiment	Experiment control	NA		10	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	10650016	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176599	CNhi10607	AGT.experiment	Negative control	NA		11	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12455292	47	CAGE	ASO NC_A	HDF (neonata)
DRR176598	CNhi10607	ACG.experiment	Targeted	NA		10	RP11-660L16.2	Homo sapiens	FANTOM6 project	PRJDB7993	13784084	47	CAGE	ASO G0254682_06	HDF (neonata)
DRR176597	CNhi10607	ACC.experiment	Negative control	NA		11	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13823738	47	CAGE	ASO NC_A	HDF (neonata)
DRR176596	CNhi10606	TAC.experiment	Targeted	NA		10	RP11-660L16.2	Homo sapiens	FANTOM6 project	PRJDB7993	9365743	47	CAGE	ASO G0254682_02	HDF (neonata)
DRR176531	CNhi10586	ATG.experiment	Targeted	NA		9	CTD-2319I12.4	Homo sapiens	FANTOM6 project	PRJDB7993	11254646	47	CAGE	ASO G0267248_02	HDF (neonata)

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176530	CNhi10586	AGT.experiment	Targeted	NA	9	AC007773.2	Homo sapiens	FANTOM6 project	PRJDB7993	10734435	47	CAGE	ASO G0267213 10	HDF (neonata)
DRR176529	CNhi10586	ACG.experiment	Targeted	NA	9	CTD-2162K18.5	Homo sapiens	FANTOM6 project	PRJDB7993	16065686	47	CAGE	ASO G0267254 07	HDF (neonata)
DRR176528	CNhi10586	ACC.experiment	Targeted	NA	9	AC007773.2	Homo sapiens	FANTOM6 project	PRJDB7993	11328455	47	CAGE	ASO G0267213 06	HDF (neonata)
DRR176527	CNhi10585	TAC.experiment	Targeted	NA	9	AC007773.2	Homo sapiens	FANTOM6 project	PRJDB7993	13886267	47	CAGE	ASO G0267213 03	HDF (neonata)
DRR176526	CNhi10585	GCT.experiment	Targeted	NA	9	AC007773.2	Homo sapiens	FANTOM6 project	PRJDB7993	11443613	47	CAGE	ASO G0267213 06	HDF (neonata)
DRR176525	CNhi10585	GCG.experiment	Targeted	NA	9	AC007773.2	Homo sapiens	FANTOM6 project	PRJDB7993	14293304	47	CAGE	ASO G0267213 01	HDF (neonata)
DRR176524	CNhi10585	CAC.experiment	Targeted	NA	9	RP13-20L14.6	Homo sapiens	FANTOM6 project	PRJDB7993	13906214	47	CAGE	ASO G0265458 10	HDF (neonata)
DRR176459	CNhi10577	AGT.experiment	Targeted	NA	6	RP11-357C3.3	Homo sapiens	FANTOM6 project	PRJDB7993	12766602	47	CAGE	ASO G0235437 06	HDF (neonata)
DRR176458	CNhi10577	ACG.experiment	Targeted	NA	6	RP11-693N9.2	Homo sapiens	FANTOM6 project	PRJDB7993	12342944	47	CAGE	ASO G0235505 01	HDF (neonata)
DRR176457	CNhi10577	ACC.experiment	Targeted	NA	6	RP11-357C3.3	Homo sapiens	FANTOM6 project	PRJDB7993	10234217	47	CAGE	ASO G0235437 05	HDF (neonata)
DRR176456	CNhi10544	TAC.experiment	Targeted	NA	5	FLVCR1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14941621	47	CAGE	ASO G0198468 07	HDF (neonata)
DRR176455	CNhi10544	GCT.experiment	Targeted	NA	6	RP11-2201.1.1	Homo sapiens	FANTOM6 project	PRJDB7993	7116218	47	CAGE	ASO G0233137 02	HDF (neonata)
DRR176454	CNhi10544	GCG.experiment	Targeted	NA	6	RP11-2201.1.1	Homo sapiens	FANTOM6 project	PRJDB7993	10388734	47	CAGE	ASO G0233137 02	HDF (neonata)
DRR176453	CNhi10544	CAC.experiment	Targeted	NA	6	AC012146.7	Homo sapiens	FANTOM6 project	PRJDB7993	12177645	47	CAGE	ASO G0234327 07	HDF (neonata)
DRR176452	CNhi10544	ATG.experiment	Targeted	NA	5	CLDN22	Homo sapiens	FANTOM6 project	PRJDB7993	12219152	47	CAGE	ASO C0008202 04	HDF (neonata)
DRR176387	CNhi10536	ACC.experiment	Targeted	NA	6	RP11-42238.1	Homo sapiens	FANTOM6 project	PRJDB7993	10945743	47	CAGE	ASO G0233621 05	HDF (neonata)
DRR176386	CNhi10535	TAC.experiment	Targeted	NA	6	RP11-2201.1.1	Homo sapiens	FANTOM6 project	PRJDB7993	18809659	47	CAGE	ASO G0233137 04	HDF (neonata)
DRR176385	CNhi10535	GCT.experiment	Targeted	NA	5	RP11-296O14.3	Homo sapiens	FANTOM6 project	PRJDB7993	18022018	47	CAGE	ASO G0203739 10	HDF (neonata)
DRR176384	CNhi10535	GCG.experiment	Targeted	NA	5	CLDN22	Homo sapiens	FANTOM6 project	PRJDB7993	12697575	47	CAGE	ASO C0008202 04	HDF (neonata)
DRR176383	CNhi10535	CAC.experiment	Targeted	NA	6	RP11-2201.1.1	Homo sapiens	FANTOM6 project	PRJDB7993	14539897	47	CAGE	ASO G0233137 05	HDF (neonata)
DRR176382	CNhi10535	ATG.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12941415	47	CAGE	ASO NC A	HDF (neonata)
DRR176381	CNhi10535	AGT.experiment	Targeted	NA	6	MAGI2-AS3	Homo sapiens	FANTOM6 project	PRJDB7993	13764596	47	CAGE	ASO G0234456 01	HDF (neonata)
DRR176380	CNhi10535	ACG.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14618235	47	CAGE	ASO NC A	HDF (neonata)
DRR176315	CNhi10526	TAC.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	17536993	47	CAGE	ASO NC A	HDF (neonata)
DRR176314	CNhi10526	GCT.experiment	Targeted	NA	4	RP11-527D7.1	Homo sapiens	FANTOM6 project	PRJDB7993	20919134	47	CAGE	ASO G0225554 02	HDF (neonata)
DRR176313	CNhi10526	GCG.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11868128	47	CAGE	ASO NC A	HDF (neonata)
DRR176312	CNhi10526	CAC.experiment	Targeted	NA	4	RP1-152L7.5	Homo sapiens	FANTOM6 project	PRJDB7993	16181406	47	CAGE	ASO G0216775 03	HDF (neonata)
DRR176311	CNhi10526	ATG.experiment	Targeted	NA	4	BOLA3-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	17261984	47	CAGE	ASO G0225439 06	HDF (neonata)
DRR176310	CNhi10526	AGT.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	18234464	47	CAGE	ASO NC A	HDF (neonata)
DRR176309	CNhi10526	ACG.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15439481	47	CAGE	ASO NC A	HDF (neonata)
DRR176308	CNhi10526	ACC.experiment	Targeted	NA	4	RP11-139H15.1	Homo sapiens	FANTOM6 project	PRJDB7993	15294128	47	CAGE	ASO G0225973 10	HDF (neonata)
DRR176243	CNhi10494	GCG.experiment	Targeted	NA	3	LINC00963	Homo sapiens	FANTOM6 project	PRJDB7993	11684910	47	CAGE	ASO G0204054 10	HDF (neonata)
DRR176242	CNhi10494	CAC.experiment	Targeted	NA	3	ZNF131	Homo sapiens	FANTOM6 project	PRJDB7993	11244694	47	CAGE	ASO C0008543 05	HDF (neonata)
DRR176241	CNhi10494	ATG.experiment	Targeted	NA	3	LINC00886	Homo sapiens	FANTOM6 project	PRJDB7993	16564637	47	CAGE	ASO G0240875 02	HDF (neonata)
DRR176240	CNhi10494	AGT.experiment	Targeted	NA	3	AC091729.9	Homo sapiens	FANTOM6 project	PRJDB7993	14372407	47	CAGE	ASO G0229043 07	HDF (neonata)
DRR176239	CNhi10494	ACG.experiment	Targeted	NA	3	FBXL19-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13696451	47	CAGE	ASO G0208852 02	HDF (neonata)
DRR176238	CNhi10494	ACC.experiment	Targeted	NA	3	AC108488.3	Homo sapiens	FANTOM6 project	PRJDB7993	11262911	47	CAGE	ASO G0234171 09	HDF (neonata)
DRR176237	CNhi10493	TAC.experiment	Targeted	NA	3	LINC00963	Homo sapiens	FANTOM6 project	PRJDB7993	16623164	47	CAGE	ASO G0204054 01	HDF (neonata)
DRR176236	CNhi10493	GCT.experiment	Targeted	NA	3	TTN-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	15248522	47	CAGE	ASO G0237298 01	HDF (neonata)
DRR176171	CNhi10477	CAC.experiment	Targeted	NA	2	CATG00000108369.1	Homo sapiens	FANTOM6 project	PRJDB7993	19517145	47	CAGE	ASO C014559 06	HDF (neonata)
DRR176170	CNhi10477	ATG.experiment	Targeted	NA	2	RP11-473M20.14	Homo sapiens	FANTOM6 project	PRJDB7993	18814708	47	CAGE	ASO G0263072 02	HDF (neonata)
DRR176169	CNhi10477	AGT.experiment	Targeted	NA	2	A1BG-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	16139368	47	CAGE	ASO G0268895 01	HDF (neonata)
DRR176168	CNhi10477	ACC.experiment	Targeted	NA	2	TRAM2-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14223437	47	CAGE	ASO G0225791 05	HDF (neonata)
DRR176167	CNhi10477	ACG.experiment	Targeted	NA	2	RP11-458D21.1	Homo sapiens	FANTOM6 project	PRJDB7993	19592447	47	CAGE	ASO G0233396 02	HDF (neonata)
DRR176166	CNhi10476	TAC.experiment	Targeted	NA	2	RP11-195F19.9	Homo sapiens	FANTOM6 project	PRJDB7993	19249934	47	CAGE	ASO G0230074 05	HDF (neonata)
DRR176165	CNhi10476	GCT.experiment	Targeted	NA	2	CATG00000089639.1	Homo sapiens	FANTOM6 project	PRJDB7993	18054700	47	CAGE	ASO G0223811 06	HDF (neonata)
DRR176164	CNhi10476	GCG.experiment	Negative control	NA	2	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13897619	47	CAGE	ASO NC A	HDF (neonata)
DRR176099	CNhi10468	AGT.experiment	Targeted	NA	2	RP11-473M20.14	Homo sapiens	FANTOM6 project	PRJDB7993	20087658	47	CAGE	ASO G0263072 01	HDF (neonata)
DRR176098	CNhi10468	ACG.experiment	Targeted	NA	2	AC005592.2	Homo sapiens	FANTOM6 project	PRJDB7993	17257014	47	CAGE	ASO G0231185 02	HDF (neonata)
DRR176097	CNhi10468	ACC.experiment	Targeted	NA	2	RP11-65J3.1	Homo sapiens	FANTOM6 project	PRJDB7993	19886678	47	CAGE	ASO G0233901 05	HDF (neonata)
DRR176096	CNhi10467	TAC.experiment	Targeted	NA	2	AC005592.2	Homo sapiens	FANTOM6 project	PRJDB7993	26105456	47	CAGE	ASO G0231185 02	HDF (neonata)
DRR177123	CNhi10837	ACG.experiment	Targeted	NA	21	RP11-398K22.12	Homo sapiens	FANTOM6 project	PRJDB7993	15907282	47	CAGE	ASO G0229852 AD 01	HDF (neonata)
DRR177122	CNhi10837	ACC.experiment	Targeted	NA	18	ZNF131	Homo sapiens	FANTOM6 project	PRJDB7993	14641374	47	CAGE	ASO C0008543 AD 10	HDF (neonata)
DRR177121	CNhi10836	TAC.experiment	Targeted	NA	19	NNT-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	19425777	47	CAGE	ASO G0248092 AD 10	HDF (neonata)
DRR177120	CNhi10836	GCT.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13591010	47	CAGE	ASO Untreat	HDF (neonata)
DRR177119	CNhi10836	GCG.experiment	Targeted	NA	21	AC091729.9	Homo sapiens	FANTOM6 project	PRJDB7993	13478630	47	CAGE	ASO G0229043 AD 04	HDF (neonata)
DRR177118	CNhi10836	CAC.experiment	Targeted	NA	19	NR2F1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13208980	47	CAGE	ASO G0237187 AD 01	HDF (neonata)
DRR177117	CNhi10836	ATG.experiment	Experiment control	NA	18	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	17019114	47	CAGE	ASO MALAT1	HDF (neonata)
DRR177116	CNhi10836	AGT.experiment	Targeted	NA	19	RP11-10K16.1	Homo sapiens	FANTOM6 project	PRJDB7993	16072090	47	CAGE	ASO G0245213 AD 04	HDF (neonata)
DRR177051	CNhi10820	ACC.experiment	Targeted	NA	20	SNRPD3	Homo sapiens	FANTOM6 project	PRJDB7993	18018036	47	CAGE	siRNA_s13234	HDF (neonata)
DRR177050	CNhi10819	TAC.experiment	Targeted	NA	20	EXOSC10	Homo sapiens	FANTOM6 project	PRJDB7993	21196719	47	CAGE	siRNA_s10737	HDF (neonata)
DRR177049	CNhi10819	GCT.experiment	Targeted	NA	20	EXOSC10	Homo sapiens	FANTOM6 project	PRJDB7993	11648131	47	CAGE	siRNA_s10738	HDF (neonata)
DRR177048	CNhi10819	GCG.experiment	Targeted	NA	21	CHAF1B	Homo sapiens	FANTOM6 project	PRJDB7993	12223552	47	CAGE	siRNA_s15706	HDF (neonata)
DRR177047	CNhi10819	CAC.experiment	Targeted	NA	21	CHAF1A	Homo sapiens	FANTOM6 project	PRJDB7993	12260612	47	CAGE	siRNA_s19500	HDF (neonata)
DRR177046	CNhi10819	ATG.experiment	Targeted	NA	21	CHAF1B	Homo sapiens	FANTOM6 project	PRJDB7993	25113885	47	CAGE	siRNA_s15705	HDF (neonata)
DRR177045	CNhi10819	AGT.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15330368	47	CAGE	siRNA_NC 1	HDF (neonata)
DRR177044	CNhi10819	ACG.experiment	Targeted	NA	20	EXOSC3	Homo sapiens	FANTOM6 project	PRJDB7993	17316875	47	CAGE	siRNA_s27231	HDF (neonata)
DRR176043	CNhi10459	ACC.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13965430	47	CAGE	ASO NC A	HDF (neonata)
DRR176042	CNhi10458	TAC.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16122752	47	CAGE	ASO NC A	HDF (neonata)
DRR176041	CNhi10458	GCT.experiment	Targeted	NA	1	RP3-510D11.2	Homo sapiens	FANTOM6 project	PRJDB7993	19033915	47	CAGE	ASO G0234546 02	HDF (neonata)
DRR176040	CNhi10458	GCG.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12105500	47	CAGE	ASO NC A	HDF (neonata)
DRR176039	CNhi10458	CAC.experiment	Targeted	NA	1	LINC00702	Homo sapiens	FANTOM6 project	PRJDB7993	14141968	47	CAGE	ASO G0233117 02	HDF (neonata)
DRR176038	CNhi10458	ATG.experiment	Targeted	NA	1	AC017048.3	Homo sapiens	FANTOM6 project	PRJDB7993	16795461	47	CAGE	ASO G0163364 06	HDF (neonata)
DRR176037	CNhi10458	AGT.experiment	Targeted	NA	1	PTCHD3P1	Homo sapiens	FANTOM6 project	PRJDB7993	19406130	47	CAGE	ASO G0224597 06	HDF (neonata)
DRR176036	CNhi10458	ACG.experiment	Experiment control	NA	1	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	22110185	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176979	CNhi10810	TAC.experiment	Targeted	NA	21	HOTAIR	Homo sapiens	FANTOM6 project	PRJDB7993	23995884	47	CAGE	ASO G0228630 POS 06	HDF (neonata)
DRR176978	CNhi10810	GCT.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12677649	47	CAGE	ASO NC B	HDF (neonata)
DRR176977	CNhi10810	CAC.experiment	Targeted	NA	20	CDKN2B-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11160289	47	CAGE	siRNA_r509345	HDF (neonata)
DRR176976	CNhi10810	ATG.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	21539518	47	CAGE	ASO Untreat	HDF (neonata)
DRR176975	CNhi10810	AGT.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12667815	47	CAGE	siRNA_NC 1	HDF (neonata)
DRR176974	CNhi10810	ACG.experiment	Targeted	NA	21	HOTAIR	Homo sapiens	FANTOM6 project	PRJDB7993	21106543	47	CAGE	ASO G0228630 POS 07	HDF (neonata)

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176973	CNhi10810	ACC.experiment	Targeted	NA	21	HOTAIR	Homo sapiens	FANTOM6 project	PRJDB7993	19852143	47	CAGE	siRNA_ADAAYU4	HDF (neonata)
DRR176972	CNhi10809	TAC.experiment	Targeted	NA	22	LINC00862	Homo sapiens	FANTOM6 project	PRJDB7993	18042294	47	CAGE	sgRNA_blasti5_LINC00862	HDF (neonata)
DRR176965	CNhi10467	GCT.experiment	Targeted	NA	2	CATG00000089639.1	Homo sapiens	FANTOM6 project	PRJDB7993	14925711	47	CAGE	ASO_G0223811_04	HDF (neonata)
DRR176964	CNhi10467	GCG.experiment	Negative control	NA	2	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13869436	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176963	CNhi10467	CAC.experiment	Targeted	NA	2	RP11-398K22.12	Homo sapiens	FANTOM6 project	PRJDB7993	16493655	47	CAGE	ASO_G0229852_05	HDF (neonata)
DRR176962	CNhi10467	ATG.experiment	Targeted	NA	2	NEAT1_1	Homo sapiens	FANTOM6 project	PRJDB7993	17532467	47	CAGE	ASO_NEAT1_1	HDF (neonata)
DRR177099	CNhi10834	ACG.experiment	Targeted	NA	19	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	17620895	47	CAGE	ASO_G0226950_04	HDF (neonata)
DRR177098	CNhi10834	ACC.experiment	Targeted	NA	18	AHR	Homo sapiens	FANTOM6 project	PRJDB7993	19749552	47	CAGE	ASO_C0011632_AD_01	HDF (neonata)
DRR177097	CNhi10833	TAC.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	24301622	47	CAGE	ASO_Lipo	HDF (neonata)
DRR177096	CNhi10833	GCT.experiment	Targeted	NA	19	MLKL1-AS2	Homo sapiens	FANTOM6 project	PRJDB7993	12627254	47	CAGE	ASO_G0236753_AD_01	HDF (neonata)
DRR177095	CNhi10833	CAC.experiment	Targeted	NA	19	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	14096731	47	CAGE	ASO_G0226950_04	HDF (neonata)
DRR177094	CNhi10833	ATG.experiment	Targeted	NA	19	US2111.14	Homo sapiens	FANTOM6 project	PRJDB7993	20193999	47	CAGE	ASO_G0232725_AD_10	HDF (neonata)
DRR177093	CNhi10833	AGT.experiment	Targeted	NA	19	RP11-539L10.3	Homo sapiens	FANTOM6 project	PRJDB7993	13729631	47	CAGE	ASO_G0251580_AD_01	HDF (neonata)
DRR177092	CNhi10833	ACG.experiment	Targeted	NA	18	CATG00000095991.1	Homo sapiens	FANTOM6 project	PRJDB7993	20012008	47	CAGE	ASO_C0012015_AD_02	HDF (neonata)
DRR177027	CNhi10817	ACC.experiment	Targeted	NA	20	CDKN2B-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	19349488	47	CAGE	ASO_G0240498_POS_04	HDF (neonata)
DRR177026	CNhi10816	TAC.experiment	Targeted	NA	20	PRRX1	Homo sapiens	FANTOM6 project	PRJDB7993	16908584	47	CAGE	ASO_G0116132_POS_09	HDF (neonata)
DRR177025	CNhi10816	GCT.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14536073	47	CAGE	ASO_NC_A	HDF (neonata)
DRR177024	CNhi10816	GCG.experiment	Targeted	NA	21	SOC52-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11668511	47	CAGE	ASO_G0246985_AD_04	HDF (neonata)
DRR177023	CNhi10816	CAC.experiment	Negative control	NA	13	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13698916	47	CAGE	ASO_NC_B	HDF (neonata)
DRR177022	CNhi10816	AGT.experiment	Targeted	NA	21	RP11-527D7.1	Homo sapiens	FANTOM6 project	PRJDB7993	15472147	47	CAGE	ASO_G0225554_AD_06	HDF (neonata)
DRR177021	CNhi10816	ACG.experiment	Targeted	NA	20	EMC7	Homo sapiens	FANTOM6 project	PRJDB7993	17175331	47	CAGE	siRNA_s32293	HDF (neonata)
DRR177020	CNhi10816	ACC.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	19313579	47	CAGE	siRNA_NC_2	HDF (neonata)
DRR176019	CNhi10456	ACC.experiment	Targeted	NA	1	TUG1	Homo sapiens	FANTOM6 project	PRJDB7993	18721025	47	CAGE	ASO_G0253352_01	HDF (neonata)
DRR176018	CNhi10455	TAC.experiment	Targeted	NA	1	AC017048.4	Homo sapiens	FANTOM6 project	PRJDB7993	20519020	47	CAGE	ASO_G0224577_05	HDF (neonata)
DRR176017	CNhi10455	GCT.experiment	Targeted	NA	1	RP11-834C11.4	Homo sapiens	FANTOM6 project	PRJDB7993	15467122	47	CAGE	ASO_G0250742_01	HDF (neonata)
DRR176016	CNhi10455	GCG.experiment	Targeted	NA	1	LINC00938	Homo sapiens	FANTOM6 project	PRJDB7993	13337520	47	CAGE	ASO_G0273015_05	HDF (neonata)
DRR176015	CNhi10455	CAC.experiment	Targeted	NA	1	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	13079333	47	CAGE	ASO_G0230590_02	HDF (neonata)
DRR176014	CNhi10455	AGT.experiment	Targeted	NA	1	EMX2OS	Homo sapiens	FANTOM6 project	PRJDB7993	18755851	47	CAGE	ASO_G0229847_03	HDF (neonata)
DRR176013	CNhi10455	ACG.experiment	Targeted	NA	1	SNHG18	Homo sapiens	FANTOM6 project	PRJDB7993	14574960	47	CAGE	ASO_G0250786_04	HDF (neonata)
DRR176012	CNhi10455	ACC.experiment	Targeted	NA	1	FGD5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14714264	47	CAGE	ASO_G0225733_04	HDF (neonata)
DRR176955	CNhi10807	GCT.experiment	Targeted	NA	22	LINC00630	Homo sapiens	FANTOM6 project	PRJDB7993	11433567	47	CAGE	sgRNA_blasti5_LINC00630	HDF (neonata)
DRR176954	CNhi10807	GCG.experiment	Negative control	NA	22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11670939	47	CAGE	sgRNA_NC_gRNA1	HDF (neonata)
DRR176953	CNhi10807	CAC.experiment	Negative control	NA	22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15710645	47	CAGE	sgRNA_dCas9	HDF (neonata)
DRR176952	CNhi10807	ATG.experiment	Negative control	NA	22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11254028	47	CAGE	sgRNA_NC_gRNA2	HDF (neonata)
DRR176951	CNhi10807	AGT.experiment	Negative control	NA	22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	17996960	47	CAGE	sgRNA_blasti5 control	HDF (neonata)
DRR176950	CNhi10807	ACG.experiment	Targeted	NA	22	LINC00630	Homo sapiens	FANTOM6 project	PRJDB7993	15086601	47	CAGE	sgRNA_blasti5_LINC00630	HDF (neonata)
DRR176949	CNhi10807	ACC.experiment	Negative control	NA	22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	17414905	47	CAGE	sgRNA_NC	HDF (neonata)
DRR176948	CNhi10805	TAC.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	18654719	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176883	CNhi10789	CAC.experiment	Negative control	NA	16	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14216151	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176882	CNhi10789	ATG.experiment	Negative control	NA	15	NA	Homo sapiens	FANTOM6 project	PRJDB7993	23355688	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176881	CNhi10789	AGT.experiment	Targeted	NA	16	CTD-231912.4	Homo sapiens	FANTOM6 project	PRJDB7993	15157624	47	CAGE	ASO_G0267248_AD_07	HDF (neonata)
DRR176880	CNhi10789	ACG.experiment	Negative control	NA	15	NA	Homo sapiens	FANTOM6 project	PRJDB7993	20850749	47	CAGE	ASO_Untreat	HDF (neonata)
DRR176879	CNhi10789	ACC.experiment	Targeted	NA	15	LINC00674	Homo sapiens	FANTOM6 project	PRJDB7993	20924093	47	CAGE	ASO_G0237854_AD_02	HDF (neonata)
DRR176878	CNhi10788	CAC.experiment	Targeted	NA	15	LINC00551	Homo sapiens	FANTOM6 project	PRJDB7993	17561858	47	CAGE	ASO_G0227036_AD_07	HDF (neonata)
DRR176877	CNhi10788	GCT.experiment	Targeted	NA	15	CTD-2006C1.2	Homo sapiens	FANTOM6 project	PRJDB7993	15564975	47	CAGE	ASO_G0219665_AD_04	HDF (neonata)
DRR176876	CNhi10788	GCG.experiment	Targeted	NA	15	AC007879.7	Homo sapiens	FANTOM6 project	PRJDB7993	15281196	47	CAGE	ASO_G0229647_AD_09	HDF (neonata)
DRR176811	CNhi10766	AGT.experiment	Targeted	NA	14	RP11-553K8.5	Homo sapiens	FANTOM6 project	PRJDB7993	16942489	47	CAGE	ASO_G0261573_AD_04	HDF (neonata)
DRR176810	CNhi10766	ACG.experiment	Targeted	NA	14	CTD-2651B20.3	Homo sapiens	FANTOM6 project	PRJDB7993	16584690	47	CAGE	ASO_G0259520_AD_04	HDF (neonata)
DRR176809	CNhi10766	ACC.experiment	Targeted	NA	13	MEG3	Homo sapiens	FANTOM6 project	PRJDB7993	20060022	47	CAGE	ASO_G0214548_AD_06	HDF (neonata)
DRR176808	CNhi10765	TAC.experiment	Negative control	NA	14	NA	Homo sapiens	FANTOM6 project	PRJDB7993	21301310	47	CAGE	ASO_Untreat	HDF (neonata)
DRR176807	CNhi10765	GCT.experiment	Targeted	NA	14	CTD-2651B20.3	Homo sapiens	FANTOM6 project	PRJDB7993	14792847	47	CAGE	ASO_G0259520_AD_04	HDF (neonata)
DRR176806	CNhi10765	GCG.experiment	Targeted	NA	13	RP11-527D7.1	Homo sapiens	FANTOM6 project	PRJDB7993	11939892	47	CAGE	ASO_G0225554_AD_06	HDF (neonata)
DRR176805	CNhi10765	CAC.experiment	Negative control	NA	17	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13794124	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176804	CNhi10765	ATG.experiment	Negative control	NA	13	NA	Homo sapiens	FANTOM6 project	PRJDB7993	20044117	47	CAGE	ASO_Lipo	HDF (neonata)
DRR176739	CNhi10757	ACG.experiment	Targeted	NA	17	RP11-611E13.2	Homo sapiens	FANTOM6 project	PRJDB7993	19425424	47	CAGE	ASO_G0257815_AD_03	HDF (neonata)
DRR176738	CNhi10757	ACC.experiment	Targeted	NA	14	RP5-886K2.3	Homo sapiens	FANTOM6 project	PRJDB7993	16394804	47	CAGE	ASO_G0236810_AD_06	HDF (neonata)
DRR176737	CNhi10640	TAC.experiment	Experiment control	NA	12	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	15512281	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR176736	CNhi10640	GCT.experiment	Targeted	NA	12	ERVVK13-1	Homo sapiens	FANTOM6 project	PRJDB7993	14044030	47	CAGE	ASO_G0260565_AD_08	HDF (neonata)
DRR176735	CNhi10640	GCG.experiment	Targeted	NA	12	TP53TG1	Homo sapiens	FANTOM6 project	PRJDB7993	12132651	47	CAGE	ASO_G0182165_AD_04	HDF (neonata)
DRR176734	CNhi10640	CAC.experiment	Targeted	NA	12	TUG1	Homo sapiens	FANTOM6 project	PRJDB7993	14718914	47	CAGE	ASO_G0253352_AD_08	HDF (neonata)
DRR176733	CNhi10640	ATG.experiment	Targeted	NA	12	MEG3	Homo sapiens	FANTOM6 project	PRJDB7993	11278052	47	CAGE	ASO_G0214548_AD_09	HDF (neonata)
DRR176732	CNhi10640	AGT.experiment	Targeted	NA	12	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	11112329	47	CAGE	ASO_G0230590_AD_02	HDF (neonata)
DRR176667	CNhi10631	TAC.experiment	Targeted	NA	12	AC017048.4	Homo sapiens	FANTOM6 project	PRJDB7993	15599940	47	CAGE	ASO_G0224577_AD_05	HDF (neonata)
DRR176666	CNhi10631	GCT.experiment	Targeted	NA	12	RP11-221N13.3	Homo sapiens	FANTOM6 project	PRJDB7993	12547637	47	CAGE	ASO_G0256268_AD_07	HDF (neonata)
DRR176665	CNhi10631	GCG.experiment	Targeted	NA	12	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11841344	47	CAGE	ASO_G0246067_AD_04	HDF (neonata)
DRR176664	CNhi10631	CAC.experiment	Targeted	NA	12	PTCHD3P1	Homo sapiens	FANTOM6 project	PRJDB7993	15838681	47	CAGE	ASO_G0224597_AD_04	HDF (neonata)
DRR176663	CNhi10631	ATG.experiment	Targeted	NA	12	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	15802577	47	CAGE	ASO_G0246067_AD_06	HDF (neonata)
DRR176662	CNhi10631	AGT.experiment	Targeted	NA	12	FGD5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11566342	47	CAGE	ASO_G0225733_AD_06	HDF (neonata)
DRR176661	CNhi10631	ACG.experiment	Targeted	NA	12	DNM3OS	Homo sapiens	FANTOM6 project	PRJDB7993	13164199	47	CAGE	ASO_G0230630_AD_10	HDF (neonata)
DRR176660	CNhi10631	ACC.experiment	Targeted	NA	12	RP11-834C11.4	Homo sapiens	FANTOM6 project	PRJDB7993	16237062	47	CAGE	ASO_G0250742_AD_05	HDF (neonata)
DRR176595	CNhi10606	GCT.experiment	Targeted	NA	10	SBF2-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14311399	47	CAGE	ASO_G0246273_02	HDF (neonata)
DRR176594	CNhi10606	GCG.experiment	Targeted	NA	10	RP11-660L16.2	Homo sapiens	FANTOM6 project	PRJDB7993	10244953	47	CAGE	ASO_G0254682_02	HDF (neonata)
DRR176593	CNhi10606	CAC.experiment	Targeted	NA	10	SRP14-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	9740852	47	CAGE	ASO_G0248508_10	HDF (neonata)
DRR176592	CNhi10606	ATG.experiment	Targeted	NA	11	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	9542937	47	CAGE	ASO_G0246067_AD_04	HDF (neonata)
DRR176591	CNhi10606	AGT.experiment	Targeted	NA	10	CKMT2-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13409484	47	CAGE	ASO_G0247572_09	HDF (neonata)
DRR176590	CNhi10606	ACC.experiment	Targeted	NA	10	RP11-660L16.2	Homo sapiens	FANTOM6 project	PRJDB7993	13362869	47	CAGE	ASO_G0254682_06	HDF (neonata)
DRR176589	CNhi10606	ACC.experiment	Targeted	NA	10	CTD-3018O17.3	Homo sapiens	FANTOM6 project	PRJDB7993	10561558	47	CAGE	ASO_G0269834_02	HDF (neonata)
DRR176588	CNhi10605	TAC.experiment	Targeted	NA	11	LINC00938	Homo sapiens	FANTOM6 project	PRJDB7993	16070602	47	CAGE	ASO_G0273015_AD_10	HDF (neonata)
DRR176523	CNhi10585	ATG.experiment	Targeted	NA	9	AC007773.2	Homo sapiens	FANTOM6 project	PRJDB7993	11458291	47	CAGE	ASO_G0267213_01	HDF (neonata)
DRR176522	CNhi10585	AGT.experiment	Targeted	NA	9	RP13-20L14.6	Homo sapiens	FANTOM6 project	PRJDB7993	12006346	47	CAGE	ASO_G0265458_10	HDF (neonata)
DRR176521	CNhi10585	ACG.experiment	Targeted	NA	9	AC007773.2	Homo sapiens	FANTOM6 project	PRJDB7993	12340199	47	CAGE	ASO_G0267213_03	HDF (neonata)
DRR176520	CNhi10585	ACC.experiment	Targeted	NA	9	RP11-6N17.4	Homo sapiens	FANTOM6 project	PRJDB7993	14121901	47			

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176519	CNhi10584	TAC.experiment	Targeted	NA	9	LINC00667	Homo sapiens	FANTOM6 project	PRJDB7993	13909884	47	CAGE	ASO	G0263753	01	HDF (neonata)
DRR176518	CNhi10584	GCT.experiment	Targeted	NA	9	RP11-6N17.4	Homo sapiens	FANTOM6 project	PRJDB7993	11566175	47	CAGE	ASO	G0264920	06	HDF (neonata)
DRR176517	CNhi10584	GCG.experiment	Targeted	NA	9	RP11-95P2.1	Homo sapiens	FANTOM6 project	PRJDB7993	11983984	47	CAGE	ASO	G0262468	07	HDF (neonata)
DRR176516	CNhi10584	CAC.experiment	Targeted	NA	9	RP11-95P2.1	Homo sapiens	FANTOM6 project	PRJDB7993	17721134	47	CAGE	ASO	G0262468	04	HDF (neonata)
DRR176451	CNhi10544	AGT.experiment	Targeted	NA	5	KTNI-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	12690095	47	CAGE	ASO	G0186615	09	HDF (neonata)
DRR176450	CNhi10544	ACG.experiment	Targeted	NA	6	RP11-422J8.1	Homo sapiens	FANTOM6 project	PRJDB7993	12876174	47	CAGE	ASO	G0233621	01	HDF (neonata)
DRR176449	CNhi10544	ACG.experiment	Targeted	NA	5	CATG00000090305.1	Homo sapiens	FANTOM6 project	PRJDB7993	14512776	47	CAGE	ASO	C0011140	02	HDF (neonata)
DRR176907	CNhi10792	GCG.experiment	Targeted	NA	16	AC009005.2	Homo sapiens	FANTOM6 project	PRJDB7993	15652189	47	CAGE	ASO	G0267751	AD 02	HDF (neonata)
DRR176906	CNhi10792	CAC.experiment	Targeted	NA	15	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	14684033	47	CAGE	ASO	G0226950	04	HDF (neonata)
DRR176905	CNhi10792	ATG.experiment	Targeted	NA	16	ZNF730	Homo sapiens	FANTOM6 project	PRJDB7993	21340913	47	CAGE	ASO	G0267886	AD 04	HDF (neonata)
DRR176904	CNhi10792	AGT.experiment	Targeted	NA	15	LINC00674	Homo sapiens	FANTOM6 project	PRJDB7993	19131885	47	CAGE	ASO	G0237854	AD 04	HDF (neonata)
DRR176903	CNhi10792	ACG.experiment	Negative control	NA	15	NA	Homo sapiens	FANTOM6 project	PRJDB7993	22084137	47	CAGE	ASO	NC B		HDF (neonata)
DRR176902	CNhi10792	ACG.experiment	Targeted	NA	16	CTC-559E9.6	Homo sapiens	FANTOM6 project	PRJDB7993	22040064	47	CAGE	ASO	G0267419	AD 06	HDF (neonata)
DRR176901	CNhi10791	TAC.experiment	Targeted	NA	15	PITPNA-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	20556065	47	CAGE	ASO	G0236618	AD 07	HDF (neonata)
DRR176900	CNhi10791	GCT.experiment	Targeted	NA	16	RP13-20L14.6	Homo sapiens	FANTOM6 project	PRJDB7993	16491490	47	CAGE	ASO	G0265458	AD 03	HDF (neonata)
DRR176835	CNhi10783	ATG.experiment	Targeted	NA	16	CTC-559E9.6	Homo sapiens	FANTOM6 project	PRJDB7993	20653423	47	CAGE	ASO	G0267419	AD 04	HDF (neonata)
DRR176834	CNhi10783	AGT.experiment	Targeted	NA	17	AC013394.2	Homo sapiens	FANTOM6 project	PRJDB7993	16016295	47	CAGE	ASO	G0272888	AD 10	HDF (neonata)
DRR176833	CNhi10783	ACG.experiment	Targeted	NA	17	AC013394.2	Homo sapiens	FANTOM6 project	PRJDB7993	22686667	47	CAGE	ASO	G0272888	AD 10	HDF (neonata)
DRR176832	CNhi10783	CAC.experiment	Targeted	NA	16	ZNF730	Homo sapiens	FANTOM6 project	PRJDB7993	20257377	47	CAGE	ASO	G0267886	AD 04	HDF (neonata)
DRR176831	CNhi10782	TAC.experiment	Targeted	NA	17	CTD-3018017.3	Homo sapiens	FANTOM6 project	PRJDB7993	26329098	47	CAGE	ASO	G0269834	AD 06	HDF (neonata)
DRR176830	CNhi10782	GCT.experiment	Targeted	NA	15	LINC00674	Homo sapiens	FANTOM6 project	PRJDB7993	15740880	47	CAGE	ASO	G0237854	AD 02	HDF (neonata)
DRR176829	CNhi10782	GCG.experiment	Targeted	NA	16	LINC00667	Homo sapiens	FANTOM6 project	PRJDB7993	15907737	47	CAGE	ASO	G0263753	AD 01	HDF (neonata)
DRR176828	CNhi10782	CAC.experiment	Negative control	NA	15	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15353713	47	CAGE	ASO	Lipo		HDF (neonata)
DRR176763	CNhi10760	ACG.experiment	Targeted	NA	14	RP11-553K8.5	Homo sapiens	FANTOM6 project	PRJDB7993	17101439	47	CAGE	ASO	G0261573	AD 06	HDF (neonata)
DRR176762	CNhi10760	ACC.experiment	Negative control	NA	13	NA	Homo sapiens	FANTOM6 project	PRJDB7993	20527193	47	CAGE	ASO	NC B		HDF (neonata)
DRR176761	CNhi10759	TAC.experiment	Targeted	NA	14	LINC00959	Homo sapiens	FANTOM6 project	PRJDB7993	23080464	47	CAGE	ASO	G0237489	AD 01	HDF (neonata)
DRR176760	CNhi10759	GCT.experiment	Targeted	NA	14	RP11-173B14.5	Homo sapiens	FANTOM6 project	PRJDB7993	16271857	47	CAGE	ASO	G0261105	AD 05	HDF (neonata)
DRR176759	CNhi10759	GCG.experiment	Targeted	NA	14	RP11-173B14.5	Homo sapiens	FANTOM6 project	PRJDB7993	12531909	47	CAGE	ASO	G0261105	AD 05	HDF (neonata)
DRR176758	CNhi10759	CAC.experiment	Negative control	NA	14	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16338392	47	CAGE	ASO	NC B		HDF (neonata)
DRR176757	CNhi10759	ATG.experiment	Targeted	NA	13	RP11-137L10.6	Homo sapiens	FANTOM6 project	PRJDB7993	13426352	47	CAGE	ASO	G0221817	AD 08	HDF (neonata)
DRR176756	CNhi10759	AGT.experiment	Targeted	NA	14	CTD-265I820.3	Homo sapiens	FANTOM6 project	PRJDB7993	16115894	47	CAGE	ASO	G0259520	AD 08	HDF (neonata)
DRR176691	CNhi10635	ACC.experiment	Negative control	NA	12	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15891026	47	CAGE	ASO	NC A		HDF (neonata)
DRR176690	CNhi10634	TAC.experiment	Targeted	NA	12	SNHG18	Homo sapiens	FANTOM6 project	PRJDB7993	17051215	47	CAGE	ASO	G0250786	AD 10	HDF (neonata)
DRR176689	CNhi10634	GCT.experiment	Targeted	NA	12	DNM30S	Homo sapiens	FANTOM6 project	PRJDB7993	12977795	47	CAGE	ASO	G0230630	AD 10	HDF (neonata)
DRR176688	CNhi10634	GCG.experiment	Targeted	NA	12	SNHG18	Homo sapiens	FANTOM6 project	PRJDB7993	16568914	47	CAGE	ASO	G0250786	AD 10	HDF (neonata)
DRR176687	CNhi10634	CAC.experiment	Targeted	NA	12	RP11-834C11.4	Homo sapiens	FANTOM6 project	PRJDB7993	14583182	47	CAGE	ASO	G0250742	AD 05	HDF (neonata)
DRR176686	CNhi10634	ATG.experiment	Targeted	NA	12	EMX20S	Homo sapiens	FANTOM6 project	PRJDB7993	14043996	47	CAGE	ASO	G0228847	AD 04	HDF (neonata)
DRR176685	CNhi10634	AGT.experiment	Targeted	NA	12	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	12568153	47	CAGE	ASO	G0246067	AD 09	HDF (neonata)
DRR176684	CNhi10634	ACG.experiment	Negative control	NA	12	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13648288	47	CAGE	ASO	NC A		HDF (neonata)
DRR176619	CNhi10609	GCT.experiment	Targeted	NA	10	LINC00959	Homo sapiens	FANTOM6 project	PRJDB7993	12569837	47	CAGE	ASO	G0237489	09	HDF (neonata)
DRR176618	CNhi10609	GCG.experiment	Targeted	NA	10	RP11-110I1.12	Homo sapiens	FANTOM6 project	PRJDB7993	9801806	47	CAGE	ASO	G0255121	02	HDF (neonata)
DRR176617	CNhi10609	CAC.experiment	Negative control	NA	11	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13965835	47	CAGE	ASO	NC A		HDF (neonata)
DRR176616	CNhi10609	ATG.experiment	Targeted	NA	10	RP11-110I1.12	Homo sapiens	FANTOM6 project	PRJDB7993	9490839	47	CAGE	ASO	G0255121	03	HDF (neonata)
DRR176615	CNhi10609	AGT.experiment	Targeted	NA	10	ZNF730	Homo sapiens	FANTOM6 project	PRJDB7993	11403080	47	CAGE	ASO	G0267886	03	HDF (neonata)
DRR176614	CNhi10609	ACG.experiment	Targeted	NA	10	NR2F1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	9651859	47	CAGE	ASO	G0237187	01	HDF (neonata)
DRR176613	CNhi10609	ACC.experiment	Targeted	NA	10	CTC-228N24.3	Homo sapiens	FANTOM6 project	PRJDB7993	12804293	47	CAGE	ASO	G0245937	01	HDF (neonata)
DRR176612	CNhi10608	TAC.experiment	Targeted	NA	11	LINC00938	Homo sapiens	FANTOM6 project	PRJDB7993	13123402	47	CAGE	ASO	G0273015	05	HDF (neonata)
DRR176547	CNhi10588	GCG.experiment	Negative control	NA	9	NA	Homo sapiens	FANTOM6 project	PRJDB7993	9928893	47	CAGE	ASO	NC A		HDF (neonata)
DRR176546	CNhi10588	CAC.experiment	Negative control	NA	9	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12751916	47	CAGE	ASO	NC A		HDF (neonata)
DRR176545	CNhi10588	ATG.experiment	Negative control	NA	9	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12309647	47	CAGE	ASO	NC A		HDF (neonata)
DRR176544	CNhi10588	AGT.experiment	Negative control	NA	9	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13338104	47	CAGE	ASO	NC A		HDF (neonata)
DRR176543	CNhi10588	ACG.experiment	Negative control	NA	9	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11546490	47	CAGE	ASO	NC A		HDF (neonata)
DRR176542	CNhi10588	ACC.experiment	Negative control	NA	9	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12921391	47	CAGE	ASO	NC A		HDF (neonata)
DRR176541	CNhi10587	TAC.experiment	Experiment control	NA	9	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	12340215	47	CAGE	ASO	MALAT1		HDF (neonata)
DRR176540	CNhi10587	GCG.experiment	Targeted	NA	9	AC009005.2	Homo sapiens	FANTOM6 project	PRJDB7993	13579855	47	CAGE	ASO	G0267751	07	HDF (neonata)
DRR176475	CNhi10579	AGT.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12866319	47	CAGE	ASO	NC A		HDF (neonata)
DRR176474	CNhi10579	ACG.experiment	Targeted	NA	6	RP11-245M24.1	Homo sapiens	FANTOM6 project	PRJDB7993	10659729	47	CAGE	ASO	G0229563	04	HDF (neonata)
DRR176473	CNhi10579	ACC.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13086624	47	CAGE	ASO	NC A		HDF (neonata)
DRR176472	CNhi10578	TAC.experiment	Experiment control	NA	6	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	14184428	47	CAGE	ASO	MALAT1		HDF (neonata)
DRR176471	CNhi10578	GCT.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13324765	47	CAGE	ASO	NC A		HDF (neonata)
DRR176470	CNhi10578	GCG.experiment	Targeted	NA	6	PITPNA-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	15149146	47	CAGE	ASO	G0236618	07	HDF (neonata)
DRR176469	CNhi10578	CAC.experiment	Targeted	NA	6	LINC00263	Homo sapiens	FANTOM6 project	PRJDB7993	13753762	47	CAGE	ASO	G0235823	03	HDF (neonata)
DRR176468	CNhi10578	ATG.experiment	Experiment control	NA	6	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	13082652	47	CAGE	ASO	MALAT1		HDF (neonata)
DRR176403	CNhi10538	ACG.experiment	Targeted	NA	6	ZNF37BP	Homo sapiens	FANTOM6 project	PRJDB7993	12459693	47	CAGE	ASO	G0234420	01	HDF (neonata)
DRR176402	CNhi10538	CAC.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12329878	47	CAGE	ASO	NC A		HDF (neonata)
DRR176401	CNhi10537	TAC.experiment	Experiment control	NA	5	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	15048426	47	CAGE	ASO	MALAT1		HDF (neonata)
DRR176400	CNhi10537	GCT.experiment	Targeted	NA	6	RP11-422J8.1	Homo sapiens	FANTOM6 project	PRJDB7993	17333569	47	CAGE	ASO	G0233621	05	HDF (neonata)
DRR176399	CNhi10537	GCG.experiment	Experiment control	NA	6	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	9629498	47	CAGE	ASO	MALAT1		HDF (neonata)
DRR176398	CNhi10537	CAC.experiment	Reference	NA	0	NA	Homo sapiens	FANTOM6 project	PRJDB7993	9434856	47	CAGE	NA			HDF (neonata)
DRR176397	CNhi10537	ATG.experiment	Targeted	NA	6	AC012146.7	Homo sapiens	FANTOM6 project	PRJDB7993	13874137	47	CAGE	ASO	G0234327	01	HDF (neonata)
DRR176396	CNhi10537	AGT.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14349519	47	CAGE	ASO	NC A		HDF (neonata)
DRR176331	CNhi10528	TAC.experiment	Targeted	NA	5	WDFY3-AS2	Homo sapiens	FANTOM6 project	PRJDB7993	14609095	47	CAGE	ASO	G0180769	03	HDF (neonata)
DRR176330	CNhi10528	GCT.experiment	Targeted	NA	4	SUGT1P	Homo sapiens	FANTOM6 project	PRJDB7993	14786932	47	CAGE	ASO	G0226823	08	HDF (neonata)
DRR176329	CNhi10528	GCG.experiment	Targeted	NA	4	MF12-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13247145	47	CAGE	ASO	G0228109	06	HDF (neonata)
DRR176328	CNhi10528	CAC.experiment	Targeted	NA	4	RP5-1103G7.4	Homo sapiens	FANTOM6 project	PRJDB7993	17624981	47	CAGE	ASO	G0225377	05	HDF (neonata)
DRR176327	CNhi10528	ATG.experiment	Targeted	NA	4	SUGT1P	Homo sapiens	FANTOM6 project	PRJDB7993	15491267	47	CAGE	ASO	G0226823	03	HDF (neonata)
DRR176326	CNhi10528	AGT.experiment	Targeted	NA	4	RP11-545E17.3	Homo sapiens	FANTOM6 project	PRJDB7993	13621681	47	CAGE	ASO	G0223478	03	HDF (neonata)
DRR176325	CNhi10528	ACG.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12140679	47	CAGE	ASO	NC A		HDF (neonata)</

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176324	CNhi10528	ACC.experiment	Targeted	NA	4	MF12-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	17637147	47	CAGE	ASO G0228109_07	HDF (neonata)
DRR176259	CNhi10496	GCT.experiment	Targeted	NA	3	RP11-38L15.3	Homo sapiens	FANTOM6 project	PRJDB7993	13811608	47	CAGE	ASO G0231187_07	HDF (neonata)
DRR176258	CNhi10496	GCG.experiment	Negative control	NA	3	NA	Homo sapiens	FANTOM6 project	PRJDB7993	10080491	47	CAGE	ASO NC_A	HDF (neonata)
DRR176257	CNhi10496	CAC.experiment	Targeted	NA	3	AC124789.1	Homo sapiens	FANTOM6 project	PRJDB7993	11915783	47	CAGE	ASO G0260833_06	HDF (neonata)
DRR176256	CNhi10496	ATG.experiment	Targeted	NA	3	LINC00963	Homo sapiens	FANTOM6 project	PRJDB7993	13757483	47	CAGE	ASO G0204054_06	HDF (neonata)
DRR176255	CNhi10496	AGT.experiment	Targeted	NA	3	TTN-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	15072718	47	CAGE	ASO G0237298_01	HDF (neonata)
DRR176254	CNhi10496	ACG.experiment	Targeted	NA	3	AC108488.3	Homo sapiens	FANTOM6 project	PRJDB7993	12892467	47	CAGE	ASO G0234171_06	HDF (neonata)
DRR176253	CNhi10496	ACC.experiment	Targeted	NA	3	CATG00000098831.1	Homo sapiens	FANTOM6 project	PRJDB7993	13083730	47	CAGE	ASO C0013510_06	HDF (neonata)
DRR176252	CNhi10495	TAC.experiment	Negative control	NA	3	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13226545	47	CAGE	ASO NC_A	HDF (neonata)
DRR176187	CNhi10487	CAC.experiment	Targeted	NA	3	LINC00856	Homo sapiens	FANTOM6 project	PRJDB7993	14765856	47	CAGE	ASO G0230417_01	HDF (neonata)
DRR176186	CNhi10487	ATG.experiment	Targeted	NA	3	LINC00862	Homo sapiens	FANTOM6 project	PRJDB7993	15986904	47	CAGE	ASO G0203721_05	HDF (neonata)
DRR176185	CNhi10487	AGT.experiment	Targeted	NA	3	AC009948.5	Homo sapiens	FANTOM6 project	PRJDB7993	17227514	47	CAGE	ASO G0223660_08	HDF (neonata)
DRR176184	CNhi10487	ACG.experiment	Targeted	NA	3	LINC00963	Homo sapiens	FANTOM6 project	PRJDB7993	14351965	47	CAGE	ASO G0204054_01	HDF (neonata)
DRR176183	CNhi10487	ACC.experiment	Targeted	NA	3	CTD-2587H24.5	Homo sapiens	FANTOM6 project	PRJDB7993	14785282	47	CAGE	ASO G0267577_03	HDF (neonata)
DRR176182	CNhi10486	CAC.experiment	Targeted	NA	3	LINC00963	Homo sapiens	FANTOM6 project	PRJDB7993	16962250	47	CAGE	ASO G0204054_06	HDF (neonata)
DRR176181	CNhi10486	GCT.experiment	Targeted	NA	3	LINC00886	Homo sapiens	FANTOM6 project	PRJDB7993	17019624	47	CAGE	ASO G0240875_10	HDF (neonata)
DRR176345	CNhi10530	GCT.experiment	Targeted	NA	4	RP1-152L7.5	Homo sapiens	FANTOM6 project	PRJDB7993	11260651	47	CAGE	ASO G0216775_01	HDF (neonata)
DRR176344	CNhi10530	CAC.experiment	Negative control	NA	4	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11601665	47	CAGE	ASO NC_A	HDF (neonata)
DRR176343	CNhi10530	ATG.experiment	Reference	NA	0	NA	Homo sapiens	FANTOM6 project	PRJDB7993	8941155	47	CAGE	NA	HDF (neonata)
DRR176342	CNhi10530	AGT.experiment	Targeted	NA	4	LL22NC03-86G7.1	Homo sapiens	FANTOM6 project	PRJDB7993	12302771	47	CAGE	ASO G0224086_03	HDF (neonata)
DRR176341	CNhi10530	ACG.experiment	Targeted	NA	5	WDFY3-AS2	Homo sapiens	FANTOM6 project	PRJDB7993	10179548	47	CAGE	ASO G0180769_08	HDF (neonata)
DRR176340	CNhi10530	ACC.experiment	Targeted	NA	4	BOLA3-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11320136	47	CAGE	ASO G0225439_06	HDF (neonata)
DRR176275	CNhi10521	GCT.experiment	Targeted	NA	5	CDKN2C	Homo sapiens	FANTOM6 project	PRJDB7993	18498761	47	CAGE	ASO C0000830_08	HDF (neonata)
DRR176274	CNhi10521	GCG.experiment	Targeted	NA	4	RP11-139H15.1	Homo sapiens	FANTOM6 project	PRJDB7993	14535781	47	CAGE	ASO G0225973_10	HDF (neonata)
DRR176273	CNhi10521	CAC.experiment	Targeted	NA	4	DNAJC27-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	20745653	47	CAGE	ASO G0224165_10	HDF (neonata)
DRR176272	CNhi10521	ATG.experiment	Reference	NA	0	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11175412	47	CAGE	NA	HDF (neonata)
DRR176271	CNhi10521	AGT.experiment	Negative control	NA	4	NA	Homo sapiens	FANTOM6 project	PRJDB7993	18154652	47	CAGE	ASO NC_A	HDF (neonata)
DRR176270	CNhi10521	ACG.experiment	Targeted	NA	4	RP1-152L7.5	Homo sapiens	FANTOM6 project	PRJDB7993	19076896	47	CAGE	ASO G0216775_03	HDF (neonata)
DRR176269	CNhi10521	ACC.experiment	Targeted	NA	5	WDFY3-AS2	Homo sapiens	FANTOM6 project	PRJDB7993	20498112	47	CAGE	ASO G0180769_03	HDF (neonata)
DRR176268	CNhi10497	TAC.experiment	Negative control	NA	3	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12936368	47	CAGE	ASO NC_A	HDF (neonata)
DRR176203	CNhi10489	GCG.experiment	Targeted	NA	3	CATG00000017883.1	Homo sapiens	FANTOM6 project	PRJDB7993	12590848	47	CAGE	ASO C0020672_01	HDF (neonata)
DRR176202	CNhi10489	CAC.experiment	Targeted	NA	3	LINC00517	Homo sapiens	FANTOM6 project	PRJDB7993	12971716	47	CAGE	ASO G0259091_01	HDF (neonata)
DRR176201	CNhi10489	AGT.experiment	Targeted	NA	3	TTN-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14079909	47	CAGE	ASO G0237298_09	HDF (neonata)
DRR176200	CNhi10489	ACG.experiment	Targeted	NA	3	LINC00862	Homo sapiens	FANTOM6 project	PRJDB7993	14969455	47	CAGE	ASO G0203721_02	HDF (neonata)
DRR176199	CNhi10489	ACC.experiment	Targeted	NA	3	RP11-500C11.3	Homo sapiens	FANTOM6 project	PRJDB7993	15965980	47	CAGE	ASO G0272269_06	HDF (neonata)
DRR176198	CNhi10488	TAC.experiment	Targeted	NA	3	CATG00000017883.1	Homo sapiens	FANTOM6 project	PRJDB7993	16398204	47	CAGE	ASO C0020672_02	HDF (neonata)
DRR176197	CNhi10488	GCT.experiment	Targeted	NA	3	CATG00000098831.1	Homo sapiens	FANTOM6 project	PRJDB7993	13970654	47	CAGE	ASO C0013510_06	HDF (neonata)
DRR176196	CNhi10488	GCG.experiment	Targeted	NA	3	AC108488.3	Homo sapiens	FANTOM6 project	PRJDB7993	11923428	47	CAGE	ASO G0234171_03	HDF (neonata)
DRR176131	CNhi10472	ATG.experiment	Targeted	NA	2	RP11-115C21.2	Homo sapiens	FANTOM6 project	PRJDB7993	23900630	47	CAGE	ASO G0246089_02	HDF (neonata)
DRR176130	CNhi10472	AGT.experiment	Targeted	NA	2	LINC00630	Homo sapiens	FANTOM6 project	PRJDB7993	17445786	47	CAGE	ASO G0223546_02	HDF (neonata)
DRR176129	CNhi10472	ACG.experiment	Targeted	NA	2	RP11-150D12.1	Homo sapiens	FANTOM6 project	PRJDB7993	15554815	47	CAGE	ASO C013368_02	HDF (neonata)
DRR176128	CNhi10472	ACC.experiment	Experiment control	NA	2	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	15970345	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176127	CNhi10471	CAC.experiment	Targeted	NA	2	AC016747.3	Homo sapiens	FANTOM6 project	PRJDB7993	20803423	47	CAGE	ASO G0212978_02	HDF (neonata)
DRR176126	CNhi10471	GCT.experiment	Targeted	NA	2	AC016747.3	Homo sapiens	FANTOM6 project	PRJDB7993	17568005	47	CAGE	ASO G0212978_03	HDF (neonata)
DRR176125	CNhi10471	GCG.experiment	Targeted	NA	2	CATG00000089639.1	Homo sapiens	FANTOM6 project	PRJDB7993	15619588	47	CAGE	ASO G0223811_04	HDF (neonata)
DRR176124	CNhi10471	CAC.experiment	Experiment control	NA	2	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	14874028	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176059	CNhi10461	ACC.experiment	Targeted	NA	1	AC017048.3	Homo sapiens	FANTOM6 project	PRJDB7993	17889338	47	CAGE	ASO G0163364_06	HDF (neonata)
DRR176058	CNhi10460	CAC.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13772480	47	CAGE	ASO NC_A	HDF (neonata)
DRR176057	CNhi10460	GCT.experiment	Targeted	NA	1	RP11-834C11.4	Homo sapiens	FANTOM6 project	PRJDB7993	16138053	47	CAGE	ASO G0250742_01	HDF (neonata)
DRR176056	CNhi10460	GCG.experiment	Targeted	NA	1	EMX2OS	Homo sapiens	FANTOM6 project	PRJDB7993	12429098	47	CAGE	ASO G0229847_01	HDF (neonata)
DRR176055	CNhi10460	ACC.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12640019	47	CAGE	ASO NC_A	HDF (neonata)
DRR176054	CNhi10460	ATG.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	18302823	47	CAGE	ASO NC_A	HDF (neonata)
DRR176053	CNhi10460	AGT.experiment	Experiment control	NA	1	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	18425767	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176052	CNhi10460	ACG.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14280177	47	CAGE	ASO NC_A	HDF (neonata)
DRR177131	CNhi10842	GCT.experiment	Targeted	NA	22	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	10228756	47	CAGE	sgRNA_TERC_gRNA1	HDF (neonata)
DRR177130	CNhi10842	ACG.experiment	Negative control	NA	22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	18432436	47	CAGE	sgRNA_NC_gRNA2	HDF (neonata)
DRR177129	CNhi10837	TAC.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	17411290	47	CAGE	ASO NC_A	HDF (neonata)
DRR177128	CNhi10837	GCT.experiment	Targeted	NA	19	RP11-33B1.1	Homo sapiens	FANTOM6 project	PRJDB7993	9047520	47	CAGE	ASO G0245958_AD_04	HDF (neonata)
DRR177127	CNhi10837	GCG.experiment	Targeted	NA	19	LINC00968	Homo sapiens	FANTOM6 project	PRJDB7993	10297635	47	CAGE	ASO G0246430_AD_03	HDF (neonata)
DRR177126	CNhi10837	CAC.experiment	Targeted	NA	21	WDFY3-AS2	Homo sapiens	FANTOM6 project	PRJDB7993	9926818	47	CAGE	ASO G0180769_AD_03	HDF (neonata)
DRR177125	CNhi10837	ATG.experiment	Targeted	NA	19	NR2F1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	39065738	47	CAGE	ASO G0237187_AD_08	HDF (neonata)
DRR177124	CNhi10837	AGT.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14732889	47	CAGE	ASO NC_A	HDF (neonata)
DRR177059	CNhi10821	ACC.experiment	Targeted	NA	21	CHAF1A	Homo sapiens	FANTOM6 project	PRJDB7993	23488651	47	CAGE	siRNA_s19501	HDF (neonata)
DRR177058	CNhi10820	CAC.experiment	Targeted	NA	20	SNRPD3	Homo sapiens	FANTOM6 project	PRJDB7993	25840381	47	CAGE	siRNA_s13235	HDF (neonata)
DRR177057	CNhi10820	GCT.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	10177958	47	CAGE	siRNA_NC_2	HDF (neonata)
DRR177056	CNhi10820	GCG.experiment	Targeted	NA	20	EXOSC10	Homo sapiens	FANTOM6 project	PRJDB7993	7771202	47	CAGE	siRNA_s10737	HDF (neonata)
DRR177055	CNhi10820	CAC.experiment	Targeted	NA	20	EXOSC3	Homo sapiens	FANTOM6 project	PRJDB7993	15989823	47	CAGE	siRNA_s27230	HDF (neonata)
DRR177054	CNhi10820	ATG.experiment	Targeted	NA	21	CHAF1B	Homo sapiens	FANTOM6 project	PRJDB7993	17445604	47	CAGE	siRNA_s15706	HDF (neonata)
DRR177053	CNhi10820	AGT.experiment	Targeted	NA	21	CHAF1A	Homo sapiens	FANTOM6 project	PRJDB7993	12525882	47	CAGE	siRNA_s19501	HDF (neonata)
DRR177052	CNhi10820	ACG.experiment	Targeted	NA	21	CHAF1B	Homo sapiens	FANTOM6 project	PRJDB7993	22844032	47	CAGE	siRNA_s15705	HDF (neonata)
DRR176987	CNhi10811	TAC.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	18414507	47	CAGE	ASO NC_A	HDF (neonata)
DRR176986	CNhi10811	GCT.experiment	Targeted	NA	13	RP11-296O14.3	Homo sapiens	FANTOM6 project	PRJDB7993	12763114	47	CAGE	ASO G0203739_AD_03	HDF (neonata)
DRR176985	CNhi10811	GCG.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14660145	47	CAGE	ASO LINC	HDF (neonata)
DRR176984	CNhi10811	CAC.experiment	Targeted	NA	20	CDKN2B-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11987491	47	CAGE	siRNA_r509345	HDF (neonata)
DRR176983	CNhi10811	ATG.experiment	Targeted	NA	13	RP11-149I23.3	Homo sapiens	FANTOM6 project	PRJDB7993	21288581	47	CAGE	ASO G0232110_AD_06	HDF (neonata)
DRR176982	CNhi10811	AGT.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13297601	47	CAGE	ASO NC_A	HDF (neonata)
DRR176981	CNhi10811	ACG.experiment	Targeted	NA	13	RP11-296O14.3	Homo sapiens	FANTOM6 project	PRJDB7993	20773733	47	CAGE	ASO G0203739_AD_03	HDF (neonata)
DRR176980	CNhi10811	ACC.experiment	Targeted	NA	21	SOC52-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	19793549	47	CAGE	ASO G0246985_AD_04	HDF (neonata)
DRR176915	CNhi10793	GCG.experiment	Negative control	NA	16	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13770257	47	CAGE	ASO NC_B	HDF (neonata)
DRR176914	CNhi10793	CAC.experiment	Negative control	NA	16	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15635383	47	CAGE	ASO NC_A	HDF (neonata)

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176913	CNhi10793	ATG.experiment	Targeted	NA	15	AC007879.7	Homo sapiens	FANTOM6 project	PRJDB7993	21584387	47	CAGE	ASO_G0229647_AD_09	HDF (neonata)
DRR176912	CNhi10793	ACG.experiment	Targeted	NA	15	PITPNA-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	16916306	47	CAGE	ASO_G0236618_AD_07	HDF (neonata)
DRR176911	CNhi10793	ACG.experiment	Negative control	NA	15	NA	Homo sapiens	FANTOM6 project	PRJDB7993	21634989	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176910	CNhi10793	ACC.experiment	Targeted	NA	17	CTD-3018017.3	Homo sapiens	FANTOM6 project	PRJDB7993	17851701	47	CAGE	ASO_G0269834_AD_02	HDF (neonata)
DRR176909	CNhi10792	ACC.experiment	Negative control	NA	16	NA	Homo sapiens	FANTOM6 project	PRJDB7993	17835216	47	CAGE	ASO_Lipo	HDF (neonata)
DRR176908	CNhi10792	GCT.experiment	Targeted	NA	15	LINC00511	Homo sapiens	FANTOM6 project	PRJDB7993	15778895	47	CAGE	ASO_G0227036_AD_07	HDF (neonata)
DRR176843	CNhi10784	ATG.experiment	Negative control	NA	16	NA	Homo sapiens	FANTOM6 project	PRJDB7993	21441206	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176842	CNhi10784	ATG.experiment	Targeted	NA	16	ZNF730	Homo sapiens	FANTOM6 project	PRJDB7993	13201480	47	CAGE	ASO_G0267886_AD_02	HDF (neonata)
DRR176841	CNhi10784	ATG.experiment	Targeted	NA	15	RP5-1103G7.4	Homo sapiens	FANTOM6 project	PRJDB7993	25845946	47	CAGE	ASO_G0225377_AD_05	HDF (neonata)
DRR176840	CNhi10784	ACC.experiment	Negative control	NA	16	NA	Homo sapiens	FANTOM6 project	PRJDB7993	23488015	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176839	CNhi10783	TAC.experiment	Targeted	NA	16	CTC-559E9.6	Homo sapiens	FANTOM6 project	PRJDB7993	20737933	47	CAGE	ASO_G0267419_AD_06	HDF (neonata)
DRR176838	CNhi10783	GCT.experiment	Targeted	NA	15	DNAJC27-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14842178	47	CAGE	ASO_G0224165_AD_07	HDF (neonata)
DRR176837	CNhi10783	GCG.experiment	Experiment control	NA	15	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	13439752	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR176836	CNhi10783	CAC.experiment	Targeted	NA	16	LINC00667	Homo sapiens	FANTOM6 project	PRJDB7993	15853984	47	CAGE	ASO_G0263753_AD_06	HDF (neonata)
DRR176771	CNhi10761	AGT.experiment	Targeted	NA	14	RP11-54A9.1	Homo sapiens	FANTOM6 project	PRJDB7993	19593326	47	CAGE	ASO_G0257219_AD_10	HDF (neonata)
DRR176770	CNhi10761	ACG.experiment	Targeted	NA	14	LINC00517	Homo sapiens	FANTOM6 project	PRJDB7993	15489618	47	CAGE	ASO_G0259091_AD_04	HDF (neonata)
DRR176769	CNhi10761	ACC.experiment	Negative control	NA	14	NA	Homo sapiens	FANTOM6 project	PRJDB7993	22978791	47	CAGE	ASO_Untreat	HDF (neonata)
DRR176768	CNhi10760	TAC.experiment	Targeted	NA	14	LINC00707	Homo sapiens	FANTOM6 project	PRJDB7993	21896095	47	CAGE	ASO_G0238266_AD_06	HDF (neonata)
DRR176767	CNhi10760	GCT.experiment	Negative control	NA	13	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16147032	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176766	CNhi10760	CAC.experiment	Negative control	NA	14	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12986162	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176765	CNhi10760	ATG.experiment	Targeted	NA	13	RP11-139H15.1	Homo sapiens	FANTOM6 project	PRJDB7993	22489696	47	CAGE	ASO_G0225973_AD_01	HDF (neonata)
DRR176764	CNhi10760	ATG.experiment	Targeted	NA	13	RP11-527D7.1	Homo sapiens	FANTOM6 project	PRJDB7993	17040254	47	CAGE	ASO_G0225554_AD_04	HDF (neonata)
DRR176699	CNhi10636	ACC.experiment	Negative control	NA	12	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14487498	47	CAGE	ASO_NC_A	HDF (neonata)
DRR177149	CNhi10867	ACG.experiment	Targeted	NA	23	XRN2	Homo sapiens	FANTOM6 project	PRJDB7993	21236619	47	CAGE	siRNA_siXRN-1	HDF (neonata)
DRR177148	CNhi10866	GCT.experiment	Negative control	NA	23	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12506223	47	CAGE	siRNA_NC_1	HDF (neonata)
DRR177083	CNhi10832	ACC.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	20204359	47	CAGE	ASO_NC_B	HDF (neonata)
DRR177082	CNhi10831	TAC.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	20034928	47	CAGE	ASO_NC_A	HDF (neonata)
DRR177081	CNhi10831	GCT.experiment	Targeted	NA	19	CTC-228N24.3	Homo sapiens	FANTOM6 project	PRJDB7993	13540341	47	CAGE	ASO_G0245937_AD_05	HDF (neonata)
DRR177080	CNhi10831	GCG.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16497346	47	CAGE	ASO_Untreat	HDF (neonata)
DRR177079	CNhi10831	CAC.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11544368	47	CAGE	ASO_NC_B	HDF (neonata)
DRR177078	CNhi10831	ATG.experiment	Targeted	NA	19	RP11-10K16.1	Homo sapiens	FANTOM6 project	PRJDB7993	16233696	47	CAGE	ASO_G0245213_AD_03	HDF (neonata)
DRR177077	CNhi10831	AGT.experiment	Targeted	NA	19	RP11-33B1.1	Homo sapiens	FANTOM6 project	PRJDB7993	14666472	47	CAGE	ASO_G0245958_AD_04	HDF (neonata)
DRR177076	CNhi10831	ACG.experiment	Targeted	NA	19	RP11-539L10.3	Homo sapiens	FANTOM6 project	PRJDB7993	18638109	47	CAGE	ASO_G0251580_AD_01	HDF (neonata)
DRR177011	CNhi10814	TAC.experiment	Targeted	NA	20	PRRX1	Homo sapiens	FANTOM6 project	PRJDB7993	16909072	47	CAGE	siRNA_si10745	HDF (neonata)
DRR177010	CNhi10814	GCT.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12846265	47	CAGE	siRNA_NC_2	HDF (neonata)
DRR177009	CNhi10814	GCG.experiment	Targeted	NA	20	CDKN2B-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10917296	47	CAGE	ASO_G0240498_POS_01	HDF (neonata)
DRR177008	CNhi10814	CAC.experiment	Negative control	NA	13	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14295845	47	CAGE	ASO_NC_A	HDF (neonata)
DRR177007	CNhi10814	ATG.experiment	Targeted	NA	21	HOTAIR	Homo sapiens	FANTOM6 project	PRJDB7993	20915537	47	CAGE	siRNA_AD70L13	HDF (neonata)
DRR177006	CNhi10814	AGT.experiment	Targeted	NA	21	LINC00707	Homo sapiens	FANTOM6 project	PRJDB7993	20716626	47	CAGE	ASO_G0238266_AD_06	HDF (neonata)
DRR177005	CNhi10814	ACG.experiment	Experiment control	NA	21	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	17873785	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR177004	CNhi10814	ACC.experiment	Targeted	NA	20	EMC7	Homo sapiens	FANTOM6 project	PRJDB7993	19004766	47	CAGE	siRNA_s32294	HDF (neonata)
DRR176003	CNhi10453	TAC.experiment	Targeted	NA	1	SNHG18	Homo sapiens	FANTOM6 project	PRJDB7993	21947660	47	CAGE	ASO_G0250786_04	HDF (neonata)
DRR176002	CNhi10453	GCT.experiment	Targeted	NA	1	RP3-510D11.2	Homo sapiens	FANTOM6 project	PRJDB7993	14274805	47	CAGE	ASO_G0234546_01	HDF (neonata)
DRR176001	CNhi10453	GCG.experiment	Targeted	NA	1	LINC00339	Homo sapiens	FANTOM6 project	PRJDB7993	13981771	47	CAGE	ASO_G0218510_04	HDF (neonata)
DRR176000	CNhi10453	CAC.experiment	Targeted	NA	1	RP3-510D11.2	Homo sapiens	FANTOM6 project	PRJDB7993	16892243	47	CAGE	ASO_G0234546_07	HDF (neonata)
DRR175999	CNhi10453	ATG.experiment	Experiment control	NA	1	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	18947020	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR175998	CNhi10453	AGT.experiment	Targeted	NA	1	RP6-99M1.2	Homo sapiens	FANTOM6 project	PRJDB7993	18889572	47	CAGE	ASO_G0270069_05	HDF (neonata)
DRR175997	CNhi10453	ACG.experiment	Targeted	NA	1	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	17264261	47	CAGE	ASO_G0270141_05	HDF (neonata)
DRR175996	CNhi10453	ACC.experiment	Experiment control	NA	1	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	17913584	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR176939	CNhi10804	GCT.experiment	Targeted	NA	21	TRAM2-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10047371	47	CAGE	ASO_G0225791_AD_05	HDF (neonata)
DRR176938	CNhi10804	GCG.experiment	Targeted	NA	19	MKLN1-AS2	Homo sapiens	FANTOM6 project	PRJDB7993	10397571	47	CAGE	ASO_G0236753_AD_01	HDF (neonata)
DRR176937	CNhi10804	CAC.experiment	Targeted	NA	19	RP11-33B1.1	Homo sapiens	FANTOM6 project	PRJDB7993	10470512	47	CAGE	ASO_G0245958_AD_07	HDF (neonata)
DRR176936	CNhi10804	ATG.experiment	Negative control	NA	18	NA	Homo sapiens	FANTOM6 project	PRJDB7993	17028180	47	CAGE	ASO_Lipo	HDF (neonata)
DRR176935	CNhi10804	AGT.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	17775061	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176934	CNhi10804	ACG.experiment	Targeted	NA	19	US2111.14	Homo sapiens	FANTOM6 project	PRJDB7993	15218248	47	CAGE	ASO_G0232725_AD_10	HDF (neonata)
DRR176933	CNhi10804	ACC.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14427754	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176932	CNhi10803	TAC.experiment	Negative control	NA	18	NA	Homo sapiens	FANTOM6 project	PRJDB7993	17236791	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176867	CNhi10787	CAC.experiment	Negative control	NA	17	NA	Homo sapiens	FANTOM6 project	PRJDB7993	8444174	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176866	CNhi10787	ATG.experiment	Targeted	NA	15	UBL7-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	25246450	47	CAGE	ASO_G0247240_AD_05	HDF (neonata)
DRR176865	CNhi10787	AGT.experiment	Targeted	NA	17	AC013394.2	Homo sapiens	FANTOM6 project	PRJDB7993	14415234	47	CAGE	ASO_G0272888_AD_07	HDF (neonata)
DRR176864	CNhi10787	ACG.experiment	Targeted	NA	15	AC009948.5	Homo sapiens	FANTOM6 project	PRJDB7993	23470712	47	CAGE	ASO_G0223960_AD_10	HDF (neonata)
DRR176863	CNhi10787	ACG.experiment	Targeted	NA	16	ZNF730	Homo sapiens	FANTOM6 project	PRJDB7993	23895922	47	CAGE	ASO_G0267886_AD_02	HDF (neonata)
DRR176862	CNhi10786	TAC.experiment	Negative control	NA	15	NA	Homo sapiens	FANTOM6 project	PRJDB7993	25938284	47	CAGE	ASO_Lipo	HDF (neonata)
DRR176861	CNhi10786	GCT.experiment	Targeted	NA	15	CTD-2006C1.2	Homo sapiens	FANTOM6 project	PRJDB7993	16489762	47	CAGE	ASO_G0219665_AD_08	HDF (neonata)
DRR176860	CNhi10786	GCG.experiment	Targeted	NA	15	UBL7-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13278053	47	CAGE	ASO_G0247240_AD_06	HDF (neonata)
DRR176795	CNhi10764	AGT.experiment	Targeted	NA	13	RP11-137L10.6	Homo sapiens	FANTOM6 project	PRJDB7993	16271822	47	CAGE	ASO_G0221817_AD_08	HDF (neonata)
DRR176794	CNhi10764	ACG.experiment	Experiment control	NA	14	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	17328544	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR176793	CNhi10764	ACC.experiment	Targeted	NA	14	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	21349220	47	CAGE	ASO_G0226950_04	HDF (neonata)
DRR176792	CNhi10763	TAC.experiment	Targeted	NA	14	RP11-173B14.5	Homo sapiens	FANTOM6 project	PRJDB7993	23834413	47	CAGE	ASO_G0261105_AD_06	HDF (neonata)
DRR176791	CNhi10763	GCT.experiment	Negative control	NA	14	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13306242	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176790	CNhi10763	GCG.experiment	Negative control	NA	14	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13165581	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176789	CNhi10763	CAC.experiment	Targeted	NA	14	SOC2-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	12266873	47	CAGE	ASO_G0246985_AD_04	HDF (neonata)
DRR176788	CNhi10763	ATG.experiment	Targeted	NA	14	RP11-54A9.1	Homo sapiens	FANTOM6 project	PRJDB7993	18886321	47	CAGE	ASO_G0257219_AD_05	HDF (neonata)
DRR176723	CNhi10639	ACG.experiment	Negative control	NA	12	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15779885	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176722	CNhi10639	ACC.experiment	Targeted	NA	12	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14930980	47	CAGE	ASO_G0246067_AD_04	HDF (neonata)
DRR176721	CNhi10638	TAC.experiment	Targeted	NA	12	RP11-221N13.3	Homo sapiens	FANTOM6 project	PRJDB7993	14550282	47	CAGE	ASO_G0256268_AD_06	HDF (neonata)
DRR176720	CNhi10638	GCT.experiment	Targeted	NA	12	RP11-221N13.3	Homo sapiens	FANTOM6 project	PRJDB7993	11922776	47	CAGE	ASO_G0256268_AD_03	HDF (neonata)
DRR176719	CNhi10638	CAC.experiment	Targeted	NA	12	TUG1	Homo sapiens	FANTOM6 project	PRJDB7993	13619034	47	CAGE	ASO_G0253352_AD_07	HDF (neonata)

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176718	CNhi10638	ATG.experiment	Targeted	NA	12	FGD5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	12265194	47	CAGE	ASO G0225733	AD	03	HDF (neonata)
DRR176717	CNhi10638	AGT.experiment	Targeted	NA	12	ERVK3-1	Homo sapiens	FANTOM6 project	PRJDB7993	13640863	47	CAGE	ASO G0142396	AD	07	HDF (neonata)
DRR176716	CNhi10638	ACG.experiment	Targeted	NA	12	FGD5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14653639	47	CAGE	ASO G0225733	AD	03	HDF (neonata)
DRR176651	CNhi10629	TAC.experiment	Targeted	NA	12	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	15227693	47	CAGE	ASO G0226950	AD	07	HDF (neonata)
DRR176650	CNhi10629	GCT.experiment	Targeted	NA	12	RP11-834C11.4	Homo sapiens	FANTOM6 project	PRJDB7993	14468082	47	CAGE	ASO G0250742	AD	04	HDF (neonata)
DRR176649	CNhi10629	GCG.experiment	Targeted	NA	12	LINC00057	Homo sapiens	FANTOM6 project	PRJDB7993	12055340	47	CAGE	ASO G0260032	AD	06	HDF (neonata)
DRR176648	CNhi10629	CAC.experiment	Targeted	NA	12	RP11-221N13.3	Homo sapiens	FANTOM6 project	PRJDB7993	13868209	47	CAGE	ASO G0256268	AD	06	HDF (neonata)
DRR176647	CNhi10629	ATG.experiment	Targeted	NA	12	RP11-417E7.1	Homo sapiens	FANTOM6 project	PRJDB7993	12588648	47	CAGE	ASO G0223485	AD	06	HDF (neonata)
DRR176646	CNhi10629	AGT.experiment	Targeted	NA	12	RP11-417E7.1	Homo sapiens	FANTOM6 project	PRJDB7993	9468284	47	CAGE	ASO G0223485	AD	04	HDF (neonata)
DRR176645	CNhi10629	ACG.experiment	Targeted	NA	12	RP11-221N13.3	Homo sapiens	FANTOM6 project	PRJDB7993	15161436	47	CAGE	ASO G0256268	AD	03	HDF (neonata)
DRR176644	CNhi10629	ACC.experiment	Targeted	NA	12	LINC00339	Homo sapiens	FANTOM6 project	PRJDB7993	19263473	47	CAGE	ASO G0218510	AD	03	HDF (neonata)
DRR176579	CNhi10604	GCT.experiment	Targeted	NA	10	AC013394.2	Homo sapiens	FANTOM6 project	PRJDB7993	16562877	47	CAGE	ASO G0272888	04		HDF (neonata)
DRR176578	CNhi10604	GCG.experiment	Targeted	NA	10	RP11-85K15.2	Homo sapiens	FANTOM6 project	PRJDB7993	10356078	47	CAGE	ASO G0258704	09		HDF (neonata)
DRR176577	CNhi10604	CAC.experiment	Targeted	NA	10	CKMT2-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	9509283	47	CAGE	ASO G0247572	08		HDF (neonata)
DRR176576	CNhi10604	AGT.experiment	Targeted	NA	11	LINC00038	Homo sapiens	FANTOM6 project	PRJDB7993	15925899	47	CAGE	ASO G0273015	AD	07	HDF (neonata)
DRR176575	CNhi10604	ACG.experiment	Targeted	NA	10	RP11-379H18.1	Homo sapiens	FANTOM6 project	PRJDB7993	14404782	47	CAGE	ASO G0271122	05		HDF (neonata)
DRR176574	CNhi10604	ACC.experiment	Targeted	NA	10	LINC00674	Homo sapiens	FANTOM6 project	PRJDB7993	14747677	47	CAGE	ASO G0237854	05		HDF (neonata)
DRR176573	CNhi10603	TAC.experiment	Targeted	NA	10	RP13-463N16.6	Homo sapiens	FANTOM6 project	PRJDB7993	14062234	47	CAGE	ASO G0242147	05		HDF (neonata)
DRR176572	CNhi10603	GCT.experiment	Targeted	NA	10	RP11-85K15.2	Homo sapiens	FANTOM6 project	PRJDB7993	12774702	47	CAGE	ASO G0258704	09		HDF (neonata)
DRR176507	CNhi10583	ATG.experiment	Targeted	NA	9	RP11-553K8.5	Homo sapiens	FANTOM6 project	PRJDB7993	11310071	47	CAGE	ASO G0261573	03		HDF (neonata)
DRR176506	CNhi10583	AGT.experiment	Targeted	NA	9	RP11-553K8.5	Homo sapiens	FANTOM6 project	PRJDB7993	21480085	47	CAGE	ASO G0261573	02		HDF (neonata)
DRR176505	CNhi10583	ACG.experiment	Targeted	NA	9	RP11-333E1.1	Homo sapiens	FANTOM6 project	PRJDB7993	9921275	47	CAGE	ASO G0261879	05		HDF (neonata)
DRR176504	CNhi10583	ACC.experiment	Targeted	NA	9	RP11-553K8.5	Homo sapiens	FANTOM6 project	PRJDB7993	10854719	47	CAGE	ASO G0261573	01		HDF (neonata)
DRR176503	CNhi10582	TAC.experiment	Targeted	NA	9	RP11-863P13.3	Homo sapiens	FANTOM6 project	PRJDB7993	12618978	47	CAGE	ASO G0261327	05		HDF (neonata)
DRR176502	CNhi10582	GCT.experiment	Targeted	NA	9	RP11-553K8.5	Homo sapiens	FANTOM6 project	PRJDB7993	14053115	47	CAGE	ASO G0261573	01		HDF (neonata)
DRR176501	CNhi10582	GCG.experiment	Targeted	NA	9	RP11-173B14.5	Homo sapiens	FANTOM6 project	PRJDB7993	14087403	47	CAGE	ASO G0261105	02		HDF (neonata)
DRR176500	CNhi10582	CAC.experiment	Targeted	NA	9	RP11-22P6.3	Homo sapiens	FANTOM6 project	PRJDB7993	14641007	47	CAGE	ASO G0260442	01		HDF (neonata)
DRR176435	CNhi10542	ACG.experiment	Experiment control	NA	6	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	12904962	47	CAGE	ASO MALAT1			HDF (neonata)
DRR176434	CNhi10542	ACC.experiment	Targeted	NA	5	AMZ2P1	Homo sapiens	FANTOM6 project	PRJDB7993	10584409	47	CAGE	ASO G0214174	02		HDF (neonata)
DRR176433	CNhi10541	TAC.experiment	Targeted	NA	6	RP11-422J8.1	Homo sapiens	FANTOM6 project	PRJDB7993	14366808	47	CAGE	ASO G0233621	01		HDF (neonata)
DRR176432	CNhi10541	GCT.experiment	Targeted	NA	5	KTNI-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13406050	47	CAGE	ASO G0186615	04		HDF (neonata)
DRR176431	CNhi10541	GCG.experiment	Targeted	NA	5	RP11-422K12.1	Homo sapiens	FANTOM6 project	PRJDB7993	11654694	47	CAGE	ASO G0204272	03		HDF (neonata)
DRR176698	CNhi10635	TAC.experiment	Targeted	NA	12	TP53TG1	Homo sapiens	FANTOM6 project	PRJDB7993	14594893	47	CAGE	ASO G0182165	AD	04	HDF (neonata)
DRR176697	CNhi10635	GCT.experiment	Targeted	NA	12	PTCHD3P1	Homo sapiens	FANTOM6 project	PRJDB7993	13684130	47	CAGE	ASO G0224597	AD	04	HDF (neonata)
DRR176696	CNhi10635	GCG.experiment	Targeted	NA	12	ERVK3-1	Homo sapiens	FANTOM6 project	PRJDB7993	14034484	47	CAGE	ASO G0142396	AD	07	HDF (neonata)
DRR176695	CNhi10635	CAC.experiment	Negative control	NA	12	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16460475	47	CAGE	ASO NC_A			HDF (neonata)
DRR176694	CNhi10635	ATG.experiment	Negative control	NA	12	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12700142	47	CAGE	ASO NC_A			HDF (neonata)
DRR176693	CNhi10635	AGT.experiment	Experiment control	NA	12	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	9257316	47	CAGE	ASO MALAT1			HDF (neonata)
DRR176692	CNhi10635	ACG.experiment	Negative control	NA	12	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13814300	47	CAGE	ASO NC_A			HDF (neonata)
DRR176627	CNhi10610	TAC.experiment	Targeted	NA	10	RP11-474B16.1	Homo sapiens	FANTOM6 project	PRJDB7993	12295758	47	CAGE	ASO G0257642	01		HDF (neonata)
DRR176626	CNhi10610	GCT.experiment	Targeted	NA	10	RP11-54A9.1	Homo sapiens	FANTOM6 project	PRJDB7993	10577790	47	CAGE	ASO G0257219	09		HDF (neonata)
DRR176625	CNhi10610	CAC.experiment	Targeted	NA	10	RP11-110I1.12	Homo sapiens	FANTOM6 project	PRJDB7993	11929568	47	CAGE	ASO G0255121	03		HDF (neonata)
DRR176624	CNhi10610	ATG.experiment	Negative control	NA	10	NA	Homo sapiens	FANTOM6 project	PRJDB7993	9628720	47	CAGE	ASO NC_A			HDF (neonata)
DRR176623	CNhi10610	AGT.experiment	Targeted	NA	10	NR2F1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10912025	47	CAGE	ASO G0237187	03		HDF (neonata)
DRR176622	CNhi10610	ACG.experiment	Targeted	NA	11	TP53TG1	Homo sapiens	FANTOM6 project	PRJDB7993	12181142	47	CAGE	ASO G0182165	AD	04	HDF (neonata)
DRR176621	CNhi10610	ACC.experiment	Targeted	NA	10	RP11-379H18.1	Homo sapiens	FANTOM6 project	PRJDB7993	11161834	47	CAGE	ASO G0271122	06		HDF (neonata)
DRR176620	CNhi10609	TAC.experiment	Targeted	NA	10	RP13-463N16.6	Homo sapiens	FANTOM6 project	PRJDB7993	12565195	47	CAGE	ASO G0242147	05		HDF (neonata)
DRR176555	CNhi10601	GCG.experiment	Targeted	NA	10	ZSCAN16-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11989818	47	CAGE	ASO G0269293	01		HDF (neonata)
DRR176554	CNhi10601	CAC.experiment	Targeted	NA	10	RP11-474B16.1	Homo sapiens	FANTOM6 project	PRJDB7993	11941654	47	CAGE	ASO G0257642	01		HDF (neonata)
DRR176553	CNhi10601	ATG.experiment	Negative control	NA	10	NA	Homo sapiens	FANTOM6 project	PRJDB7993	10197008	47	CAGE	ASO NC_A			HDF (neonata)
DRR176552	CNhi10601	AGT.experiment	Negative control	NA	10	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13723095	47	CAGE	ASO NC_A			HDF (neonata)
DRR176551	CNhi10601	ACG.experiment	Targeted	NA	10	RP11-54A9.1	Homo sapiens	FANTOM6 project	PRJDB7993	12165238	47	CAGE	ASO G0257219	09		HDF (neonata)
DRR176550	CNhi10601	ACC.experiment	Targeted	NA	10	RP6-109B7.3	Homo sapiens	FANTOM6 project	PRJDB7993	11755871	47	CAGE	ASO G0241990	01		HDF (neonata)
DRR176549	CNhi10588	TAC.experiment	Negative control	NA	9	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12656521	47	CAGE	ASO NC_A			HDF (neonata)
DRR176548	CNhi10588	GCT.experiment	Negative control	NA	9	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12217825	47	CAGE	ASO NC_A			HDF (neonata)
DRR176483	CNhi10580	AGT.experiment	Targeted	NA	6	RP4-794H19.4	Homo sapiens	FANTOM6 project	PRJDB7993	9888327	47	CAGE	ASO G0230812	10		HDF (neonata)
DRR176482	CNhi10580	ACG.experiment	Targeted	NA	6	RP11-295G20.2	Homo sapiens	FANTOM6 project	PRJDB7993	8949422	47	CAGE	ASO G0233461	03		HDF (neonata)
DRR176481	CNhi10580	ACC.experiment	Targeted	NA	6	RP4-794H19.4	Homo sapiens	FANTOM6 project	PRJDB7993	11119532	47	CAGE	ASO G0230812	05		HDF (neonata)
DRR176480	CNhi10579	TAC.experiment	Targeted	NA	6	RP11-245M24.1	Homo sapiens	FANTOM6 project	PRJDB7993	13945217	47	CAGE	ASO G0229563	04		HDF (neonata)
DRR176479	CNhi10579	GCT.experiment	Targeted	NA	6	RP4-794H19.4	Homo sapiens	FANTOM6 project	PRJDB7993	13206079	47	CAGE	ASO G0230812	05		HDF (neonata)
DRR176478	CNhi10579	GCG.experiment	Targeted	NA	6	RP11-245M24.1	Homo sapiens	FANTOM6 project	PRJDB7993	10603659	47	CAGE	ASO G0229563	02		HDF (neonata)
DRR176477	CNhi10579	CAC.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12425824	47	CAGE	ASO NC_A			HDF (neonata)
DRR176476	CNhi10579	ATG.experiment	Targeted	NA	6	RP11-245M24.1	Homo sapiens	FANTOM6 project	PRJDB7993	14068082	47	CAGE	ASO G0229563	02		HDF (neonata)
DRR176411	CNhi10539	ACG.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14217009	47	CAGE	ASO NC_A			HDF (neonata)
DRR176410	CNhi10539	ACC.experiment	Targeted	NA	6	AC092295.7	Homo sapiens	FANTOM6 project	PRJDB7993	13570676	47	CAGE	ASO G0233527	03		HDF (neonata)
DRR176409	CNhi10538	TAC.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14148704	47	CAGE	ASO NC_A			HDF (neonata)
DRR176408	CNhi10538	GCT.experiment	Experiment control	NA	5	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	16082949	47	CAGE	ASO MALAT1			HDF (neonata)
DRR176407	CNhi10538	GCG.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11512244	47	CAGE	ASO NC_A			HDF (neonata)
DRR176406	CNhi10538	CAC.experiment	Targeted	NA	6	US2111.14	Homo sapiens	FANTOM6 project	PRJDB7993	12925340	47	CAGE	ASO G0232725	05		HDF (neonata)
DRR176405	CNhi10538	ATG.experiment	Targeted	NA	5	AMZ2P1	Homo sapiens	FANTOM6 project	PRJDB7993	12916722	47	CAGE	ASO G0214174	02		HDF (neonata)
DRR176404	CNhi10538	AGT.experiment	Targeted	NA	6	AC012146.7	Homo sapiens	FANTOM6 project	PRJDB7993	13934959	47	CAGE	ASO G0234327	01		HDF (neonata)
DRR176339	CNhi10529	TAC.experiment	Targeted	NA	5	CATG00000095991.1	Homo sapiens	FANTOM6 project	PRJDB7993	15419519	47	CAGE	ASO C0012015	07		HDF (neonata)
DRR176338	CNhi10529	GCT.experiment	Targeted	NA	4	CTD-2006C1.2	Homo sapiens	FANTOM6 project	PRJDB7993	14823645	47	CAGE	ASO G0219665	08		HDF (neonata)
DRR176337	CNhi10529	GCG.experiment	Negative control	NA	4	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12564321	47	CAGE	ASO NC_A			HDF (neonata)
DRR176336	CNhi10529	CAC.experiment	Targeted	NA	4	CD99P1	Homo sapiens	FANTOM6 project	PRJDB7993	17308418	47	CAGE	ASO G0223773	07		HDF (neonata)
DRR176335	CNhi10529	ATG.experiment	Negative control	NA	4	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13666333	47	CAGE	ASO NC_A			HDF (neonata)
DRR176334	CNhi10529	AGT.experiment	Targeted	NA	4	RP5-1103G7.4	Homo sapiens	FANTOM6 project	PRJDB7993	13367028	47	CAGE	ASO G0225377	05		HDF (neonata)

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176266	CNhi10497	GCG.experiment	Targeted	NA	3	CATG00000017883.1	Homo sapiens	FANTOM6 project	PRJDB7993	8784496	47	CAGE	ASO_C0020672_07	HDF (neonata)
DRR176265	CNhi10497	CAC.experiment	Targeted	NA	3	TMCC1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	8384862	47	CAGE	ASO_G0271270_09	HDF (neonata)
DRR176264	CNhi10497	ATG.experiment	Negative control	NA	3	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12774683	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176263	CNhi10497	AGT.experiment	Targeted	NA	3	RP11-10K16.1	Homo sapiens	FANTOM6 project	PRJDB7993	12597844	47	CAGE	ASO_G0245213_04	HDF (neonata)
DRR176262	CNhi10497	ACG.experiment	Targeted	NA	3	LINC00862	Homo sapiens	FANTOM6 project	PRJDB7993	12308143	47	CAGE	ASO_G0203721_02	HDF (neonata)
DRR176261	CNhi10497	ACC.experiment	Targeted	NA	3	AC108488.3	Homo sapiens	FANTOM6 project	PRJDB7993	12643805	47	CAGE	ASO_G0234171_01	HDF (neonata)
DRR176260	CNhi10496	TAC.experiment	Targeted	NA	3	LINC00963	Homo sapiens	FANTOM6 project	PRJDB7993	14339632	47	CAGE	ASO_G0204054_10	HDF (neonata)
DRR176195	CNhi10488	CAC.experiment	Targeted	NA	3	AC124789.1	Homo sapiens	FANTOM6 project	PRJDB7993	15707409	47	CAGE	ASO_G0260833_06	HDF (neonata)
DRR176194	CNhi10488	ATG.experiment	Targeted	NA	3	RP11-38L15.3	Homo sapiens	FANTOM6 project	PRJDB7993	15567838	47	CAGE	ASO_G0231187_07	HDF (neonata)
DRR176193	CNhi10488	AGT.experiment	Targeted	NA	3	CATG00000017883.1	Homo sapiens	FANTOM6 project	PRJDB7993	15016238	47	CAGE	ASO_C0020672_02	HDF (neonata)
DRR176192	CNhi10488	ACG.experiment	Negative control	NA	3	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11010341	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176191	CNhi10488	ACC.experiment	Targeted	NA	3	AP001258.4	Homo sapiens	FANTOM6 project	PRJDB7993	17115849	47	CAGE	ASO_G0245571_06	HDF (neonata)
DRR176190	CNhi10487	TAC.experiment	Targeted	NA	3	AC004980.7	Homo sapiens	FANTOM6 project	PRJDB7993	16526913	47	CAGE	ASO_G0205485_03	HDF (neonata)
DRR176189	CNhi10487	GCT.experiment	Targeted	NA	3	RP11-38L15.3	Homo sapiens	FANTOM6 project	PRJDB7993	14278852	47	CAGE	ASO_G0231187_02	HDF (neonata)
DRR176188	CNhi10487	GCG.experiment	Targeted	NA	3	LINC00856	Homo sapiens	FANTOM6 project	PRJDB7993	12635829	47	CAGE	ASO_G0230417_02	HDF (neonata)
DRR176123	CNhi10471	ATG.experiment	Targeted	NA	2	LINC00630	Homo sapiens	FANTOM6 project	PRJDB7993	21532607	47	CAGE	ASO_G0223546_02	HDF (neonata)
DRR176122	CNhi10471	AGT.experiment	Targeted	NA	2	NEAT1_1	Homo sapiens	FANTOM6 project	PRJDB7993	22052001	47	CAGE	ASO_NEAT1_1	HDF (neonata)
DRR176121	CNhi10471	ACG.experiment	Negative control	NA	2	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13678543	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176120	CNhi10471	ACC.experiment	Targeted	NA	2	LINC00630	Homo sapiens	FANTOM6 project	PRJDB7993	21757538	47	CAGE	ASO_G0223546_01	HDF (neonata)
DRR176119	CNhi10470	TAC.experiment	Targeted	NA	2	BX322557.10	Homo sapiens	FANTOM6 project	PRJDB7993	19162591	47	CAGE	ASO_G0215447_03	HDF (neonata)
DRR176118	CNhi10470	GCT.experiment	Targeted	NA	2	AC027612.6	Homo sapiens	FANTOM6 project	PRJDB7993	17431726	47	CAGE	ASO_G0143429_05	HDF (neonata)
DRR176117	CNhi10470	GCG.experiment	Targeted	NA	2	RP11-473M20.14	Homo sapiens	FANTOM6 project	PRJDB7993	14484673	47	CAGE	ASO_G0263072_06	HDF (neonata)
DRR176116	CNhi10470	CAC.experiment	Targeted	NA	2	NEAT1_2	Homo sapiens	FANTOM6 project	PRJDB7993	15860916	47	CAGE	ASO_NEAT1_2	HDF (neonata)
DRR176051	CNhi10460	ACC.experiment	Targeted	NA	1	RP3-510D11.1	Homo sapiens	FANTOM6 project	PRJDB7993	19029363	47	CAGE	ASO_G0228526_10	HDF (neonata)
DRR176050	CNhi10459	TAC.experiment	Targeted	NA	1	TUG1	Homo sapiens	FANTOM6 project	PRJDB7993	20241534	47	CAGE	ASO_G0253352_01	HDF (neonata)
DRR176049	CNhi10459	GCT.experiment	Targeted	NA	1	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	17326968	47	CAGE	ASO_G0246067_01	HDF (neonata)
DRR176048	CNhi10459	GCG.experiment	Targeted	NA	1	RP11-417E7.1	Homo sapiens	FANTOM6 project	PRJDB7993	12758493	47	CAGE	ASO_G0223485_01	HDF (neonata)
DRR176047	CNhi10459	CAC.experiment	Targeted	NA	1	RP3-510D11.2	Homo sapiens	FANTOM6 project	PRJDB7993	15198775	47	CAGE	ASO_G0234546_07	HDF (neonata)
DRR176046	CNhi10459	ATG.experiment	Targeted	NA	1	FGD5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	18388169	47	CAGE	ASO_G0225733_03	HDF (neonata)
DRR176045	CNhi10459	AGT.experiment	Targeted	NA	1	AC017048.4	Homo sapiens	FANTOM6 project	PRJDB7993	19759990	47	CAGE	ASO_G0224577_05	HDF (neonata)
DRR176044	CNhi10459	ACC.experiment	Targeted	NA	1	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	19370324	47	CAGE	ASO_G0246067_05	HDF (neonata)
DRR175995	CNhi10452	TAC.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16181073	47	CAGE	ASO_NC_A	HDF (neonata)
DRR175994	CNhi10452	GCT.experiment	Targeted	NA	1	RP3-510D11.2	Homo sapiens	FANTOM6 project	PRJDB7993	14213039	47	CAGE	ASO_G0234546_02	HDF (neonata)
DRR175993	CNhi10452	GCG.experiment	Targeted	NA	1	LINC00883	Homo sapiens	FANTOM6 project	PRJDB7993	14642295	47	CAGE	ASO_G0243701_01	HDF (neonata)
DRR175992	CNhi10452	CAC.experiment	Experiment control	NA	1	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	16078076	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR175991	CNhi10452	ATG.experiment	Targeted	NA	1	TUG1	Homo sapiens	FANTOM6 project	PRJDB7993	19180073	47	CAGE	ASO_G0253352_02	HDF (neonata)
DRR175990	CNhi10452	AGT.experiment	Targeted	NA	1	RP3-510D11.1	Homo sapiens	FANTOM6 project	PRJDB7993	17307796	47	CAGE	ASO_G0228526_10	HDF (neonata)
DRR175989	CNhi10452	ACC.experiment	Experiment control	NA	1	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	17370390	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR175988	CNhi10452	ACG.experiment	Targeted	NA	1	AC017048.4	Homo sapiens	FANTOM6 project	PRJDB7993	21445952	47	CAGE	ASO_G0224577_02	HDF (neonata)
DRR176430	CNhi10541	CAC.experiment	Targeted	NA	5	FLVCR1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	17645667	47	CAGE	ASO_G0198468_07	HDF (neonata)
DRR176429	CNhi10541	ATG.experiment	Targeted	NA	5	RP11-622K12.1	Homo sapiens	FANTOM6 project	PRJDB7993	10890480	47	CAGE	ASO_G0204272_03	HDF (neonata)
DRR176428	CNhi10541	AGT.experiment	Targeted	NA	5	RP11-622K12.1	Homo sapiens	FANTOM6 project	PRJDB7993	12501206	47	CAGE	ASO_G0204272_04	HDF (neonata)
DRR176363	CNhi10533	ACC.experiment	Targeted	NA	6	MAGI2-AS3	Homo sapiens	FANTOM6 project	PRJDB7993	12214133	47	CAGE	ASO_G0234456_04	HDF (neonata)
DRR176362	CNhi10532	TAC.experiment	Targeted	NA	4	CD99P1	Homo sapiens	FANTOM6 project	PRJDB7993	15248670	47	CAGE	ASO_G0223773_07	HDF (neonata)
DRR176361	CNhi10532	GCT.experiment	Targeted	NA	4	RP11-137L10.6	Homo sapiens	FANTOM6 project	PRJDB7993	15881157	47	CAGE	ASO_G0221817_04	HDF (neonata)
DRR176360	CNhi10532	GCG.experiment	Targeted	NA	4	RP11-137L10.6	Homo sapiens	FANTOM6 project	PRJDB7993	11190425	47	CAGE	ASO_G0221817_04	HDF (neonata)
DRR176359	CNhi10532	CAC.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15386042	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176358	CNhi10532	ATG.experiment	Targeted	NA	4	LINC00511	Homo sapiens	FANTOM6 project	PRJDB7993	10622671	47	CAGE	ASO_G0227036_04	HDF (neonata)
DRR176357	CNhi10532	AGT.experiment	Targeted	NA	4	RP11-545E17.3	Homo sapiens	FANTOM6 project	PRJDB7993	13329693	47	CAGE	ASO_G0223478_03	HDF (neonata)
DRR176356	CNhi10532	ACC.experiment	Targeted	NA	5	AHR	Homo sapiens	FANTOM6 project	PRJDB7993	15422223	47	CAGE	ASO_C0011632_10	HDF (neonata)
DRR176291	CNhi10523	GCT.experiment	Targeted	NA	4	DNAJC27-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	16830942	47	CAGE	ASO_G0224165_10	HDF (neonata)
DRR176290	CNhi10523	GCG.experiment	Targeted	NA	4	LINC00511	Homo sapiens	FANTOM6 project	PRJDB7993	14987779	47	CAGE	ASO_G0227036_04	HDF (neonata)
DRR176289	CNhi10523	CAC.experiment	Targeted	NA	4	L122NC03-86G7.1	Homo sapiens	FANTOM6 project	PRJDB7993	23081646	47	CAGE	ASO_G0224086_04	HDF (neonata)
DRR176288	CNhi10523	ATG.experiment	Negative control	NA	4	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11890731	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176287	CNhi10523	AGT.experiment	Targeted	NA	5	CATG00000020700.1	Homo sapiens	FANTOM6 project	PRJDB7993	16260918	47	CAGE	ASO_C0020984_02	HDF (neonata)
DRR176286	CNhi10523	ACC.experiment	Targeted	NA	5	CDKN2C	Homo sapiens	FANTOM6 project	PRJDB7993	14906158	47	CAGE	ASO_C0000830_02	HDF (neonata)
DRR176285	CNhi10523	ACG.experiment	Targeted	NA	5	AHR	Homo sapiens	FANTOM6 project	PRJDB7993	17903763	47	CAGE	ASO_C0011632_03	HDF (neonata)
DRR176284	CNhi10522	TAC.experiment	Targeted	NA	4	JPX	Homo sapiens	FANTOM6 project	PRJDB7993	18088273	47	CAGE	ASO_G0225470_05	HDF (neonata)
DRR176219	CNhi10491	GCG.experiment	Targeted	NA	3	LINC00886	Homo sapiens	FANTOM6 project	PRJDB7993	13945429	47	CAGE	ASO_G0240875_01	HDF (neonata)
DRR176218	CNhi10491	CAC.experiment	Targeted	NA	3	AC108488.3	Homo sapiens	FANTOM6 project	PRJDB7993	11401092	47	CAGE	ASO_G0234171_02	HDF (neonata)
DRR176217	CNhi10491	ATG.experiment	Targeted	NA	3	CATG00000017883.1	Homo sapiens	FANTOM6 project	PRJDB7993	14802219	47	CAGE	ASO_C0020672_01	HDF (neonata)
DRR176216	CNhi10491	AGT.experiment	Targeted	NA	3	TTN-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	16412826	47	CAGE	ASO_G0237298_09	HDF (neonata)
DRR176215	CNhi10491	ACG.experiment	Targeted	NA	3	AC108488.3	Homo sapiens	FANTOM6 project	PRJDB7993	13623090	47	CAGE	ASO_G0234171_03	HDF (neonata)
DRR176214	CNhi10491	ACC.experiment	Targeted	NA	3	CTD-2587H24.5	Homo sapiens	FANTOM6 project	PRJDB7993	12080196	47	CAGE	ASO_G0267577_01	HDF (neonata)
DRR176213	CNhi10490	TAC.experiment	Targeted	NA	3	AC108488.3	Homo sapiens	FANTOM6 project	PRJDB7993	15497851	47	CAGE	ASO_G0234171_02	HDF (neonata)
DRR176212	CNhi10490	GCT.experiment	Targeted	NA	3	TMCC1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14145944	47	CAGE	ASO_G0271270_09	HDF (neonata)
DRR176147	CNhi10474	ATG.experiment	Negative control	NA	2	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16131578	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176146	CNhi10474	AGT.experiment	Targeted	NA	2	AC005592.2	Homo sapiens	FANTOM6 project	PRJDB7993	13723243	47	CAGE	ASO_G0231185_03	HDF (neonata)
DRR176145	CNhi10474	ACG.experiment	Targeted	NA	2	LINC00968	Homo sapiens	FANTOM6 project	PRJDB7993	15721364	47	CAGE	ASO_G0246430_03	HDF (neonata)
DRR176144	CNhi10474	ACC.experiment	Targeted	NA	2	AC007879.7	Homo sapiens	FANTOM6 project	PRJDB7993	13704792	47	CAGE	ASO_G0229647_06	HDF (neonata)
DRR176143	CNhi10473	TAC.experiment	Targeted	NA	2	A1BG-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	19169257	47	CAGE	ASO_G0268895_01	HDF (neonata)
DRR176142	CNhi10473	GCT.experiment	Targeted	NA	2	RP11-473M20.14	Homo sapiens	FANTOM6 project	PRJDB7993	18312725	47	CAGE	ASO_G0263072_02	HDF (neonata)
DRR176141	CNhi10473	GCG.experiment	Targeted	NA	2	CTD-3131K8.2	Homo sapiens	FANTOM6 project	PRJDB7993	13770608	47	CAGE	ASO_G0269439_09	HDF (neonata)
DRR176140	CNhi10473	CAC.experiment	Negative control	NA	2	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15661087	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176075	CNhi10463	AGT.experiment	Targeted	NA	1	LINC00702	Homo sapiens	FANTOM6 project	PRJDB7993	17043508	47	CAGE	ASO_G0233117_04	HDF (neonata)
DRR176074	CNhi10463	ACG.experiment	Targeted	NA	1	FGD5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	15200062	47	CAGE	ASO_G0225733_03	HDF (neonata)
DRR176073	CNhi10463	ACC.experiment	Targeted	NA	1	AC017048.3	Homo sapiens	FANTOM6 project	PRJDB7993	17492543	47	CAGE	ASO_G0163364_04	HDF (neonata)
DRR176072	CNhi10462	TAC.experiment	Targeted	NA	1	MEG3	Homo sapiens	FANTOM6 project	PRJDB7993	19195486	47	CAGE	ASO_G0214549_01	HDF (neonata)
DRR176071	CNhi10462	GCT.experiment	Targeted	NA	1	PTCHD3P1	Homo sapiens	FANTOM6 project	PRJDB7993	14195244	47	CAGE	ASO_G0224597_06	HDF (neonata)
DRR176070	CNhi10462	CAC.experiment	Targeted	NA	1	LINC00339	Homo sapiens	FANTOM6 project	PRJDB7993	12421822	47	CAGE	ASO_G0218510_04	HDF (neonata)
DRR176069	CNhi10462	ATG.experiment	Targeted	NA	1	LINC00702	Homo sapiens	FANTOM6 project	PRJDB7993	16219421	47	CAGE	ASO_G0233117_02	HDF (neonata)

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176068	CNhi10462	AGT.experiment	Targeted	NA	1	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	17494989	47	CAGE	ASO G0230590_05	HDF (neonata)
DRR177147	CNhi10865	TAC.experiment	Targeted	NA	23	XRN2	Homo sapiens	FANTOM6 project	PRJDB7993	22266081	47	CAGE	siRNA_sixRN-2	HDF (neonata)
DRR177146	CNhi10865	GCG.experiment	Negative control	NA	23	NA	Homo sapiens	FANTOM6 project	PRJDB7993	9212750	47	CAGE	siRNA_NC_1	HDF (neonata)
DRR177145	CNhi10865	ATG.experiment	Targeted	NA	23	XRN2	Homo sapiens	FANTOM6 project	PRJDB7993	15053272	47	CAGE	siRNA_sixRN-1	HDF (neonata)
DRR177144	CNhi10844	GCG.experiment	Targeted	NA	22	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	9995934	47	CAGE	sgRNA_FTX_gRNA3	HDF (neonata)
DRR177143	CNhi10844	CAC.experiment	Targeted	NA	22	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	12767490	47	CAGE	sgRNA_TERC_gRNA1_sgRNA4	HDF (neonata)
DRR177142	CNhi10844	ATG.experiment	Targeted	NA	22	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	20564005	47	CAGE	sgRNA_FTX_gRNA1_sgRNA4	HDF (neonata)
DRR177141	CNhi10844	AGT.experiment	Targeted	NA	22	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	15313293	47	CAGE	sgRNA_FTX_gRNA1	HDF (neonata)
DRR177140	CNhi10844	ACC.experiment	Targeted	NA	22	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	16574444	47	CAGE	sgRNA_TERC_gRNA4	HDF (neonata)
DRR177075	CNhi10831	ACC.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	20483490	47	CAGE	ASO_NC_A	HDF (neonata)
DRR177074	CNhi10830	TAC.experiment	Targeted	NA	19	CTC-228N24.3	Homo sapiens	FANTOM6 project	PRJDB7993	27162797	47	CAGE	ASO G0245937_AD_05	HDF (neonata)
DRR177073	CNhi10830	GCT.experiment	Targeted	NA	18	ZNF131	Homo sapiens	FANTOM6 project	PRJDB7993	10355362	47	CAGE	ASO C0008543_AD_10	HDF (neonata)
DRR177072	CNhi10830	GCG.experiment	Targeted	NA	19	CTC-228N24.3	Homo sapiens	FANTOM6 project	PRJDB7993	11014680	47	CAGE	ASO G0245937_AD_02	HDF (neonata)
DRR177071	CNhi10830	CAC.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13627469	47	CAGE	ASO_NC_B	HDF (neonata)
DRR177070	CNhi10830	ATG.experiment	Targeted	NA	21	JPX	Homo sapiens	FANTOM6 project	PRJDB7993	17886974	47	CAGE	ASO G0225470_AD_04	HDF (neonata)
DRR177069	CNhi10830	AGT.experiment	Targeted	NA	19	ZSCAN16-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10910014	47	CAGE	ASO G0269293_AD_10	HDF (neonata)
DRR177068	CNhi10830	ACG.experiment	Targeted	NA	19	RP11-10K16.1	Homo sapiens	FANTOM6 project	PRJDB7993	20803362	47	CAGE	ASO G0245213_AD_04	HDF (neonata)
DRR177067	CNhi10813	TAC.experiment	Targeted	NA	20	EMC7	Homo sapiens	FANTOM6 project	PRJDB7993	23272562	47	CAGE	ASO G0134153_POS_07	HDF (neonata)
DRR177002	CNhi10813	GCT.experiment	Targeted	NA	20	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	10780354	47	CAGE	siRNA_n269549	HDF (neonata)
DRR177001	CNhi10813	GCG.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11389369	47	CAGE	ASO_Untreat	HDF (neonata)
DRR177000	CNhi10813	CAC.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12916594	47	CAGE	siRNA_NC_1	HDF (neonata)
DRR176999	CNhi10813	ATG.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16964472	47	CAGE	siRNA_NC_2	HDF (neonata)
DRR176998	CNhi10813	AGT.experiment	Experiment control	NA	20	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	13265771	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR176997	CNhi10813	ACG.experiment	Targeted	NA	21	LINC00959	Homo sapiens	FANTOM6 project	PRJDB7993	17490421	47	CAGE	ASO G0237489_AD_01	HDF (neonata)
DRR176996	CNhi10813	ACG.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	19831236	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176931	CNhi10803	GCT.experiment	Targeted	NA	21	TMCC1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13589048	47	CAGE	ASO G0271270_AD_10	HDF (neonata)
DRR176930	CNhi10803	GCG.experiment	Experiment control	NA	19	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	12934259	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR176929	CNhi10803	CAC.experiment	Negative control	NA	18	NA	Homo sapiens	FANTOM6 project	PRJDB7993	9884362	47	CAGE	ASO_Untreat	HDF (neonata)
DRR176928	CNhi10803	AGT.experiment	Targeted	NA	21	ZNF674-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	15488003	47	CAGE	ASO G0230844_AD_06	HDF (neonata)
DRR176927	CNhi10803	ACG.experiment	Targeted	NA	21	TRAM2-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	17591760	47	CAGE	ASO G0225791_AD_05	HDF (neonata)
DRR176926	CNhi10803	ACC.experiment	Targeted	NA	19	RP11-10K16.1	Homo sapiens	FANTOM6 project	PRJDB7993	22654802	47	CAGE	ASO G0245213_AD_03	HDF (neonata)
DRR176925	CNhi10801	TAC.experiment	Targeted	NA	21	ZNF674-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	16429002	47	CAGE	ASO G0230844_AD_06	HDF (neonata)
DRR176924	CNhi10801	GCT.experiment	Experiment control	NA	19	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	12981015	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR176859	CNhi10786	CAC.experiment	Negative control	NA	16	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13407050	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176858	CNhi10786	ATG.experiment	Targeted	NA	15	LINC00674	Homo sapiens	FANTOM6 project	PRJDB7993	22500218	47	CAGE	ASO G0237854_AD_04	HDF (neonata)
DRR176857	CNhi10786	AGT.experiment	Targeted	NA	16	CTD-2319I12.4	Homo sapiens	FANTOM6 project	PRJDB7993	17499042	47	CAGE	ASO G0267248_AD_07	HDF (neonata)
DRR176856	CNhi10786	ACG.experiment	Negative control	NA	15	NA	Homo sapiens	FANTOM6 project	PRJDB7993	18982157	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176855	CNhi10786	ACC.experiment	Negative control	NA	16	NA	Homo sapiens	FANTOM6 project	PRJDB7993	24129021	47	CAGE	ASO_Untreat	HDF (neonata)
DRR176854	CNhi10785	TAC.experiment	Targeted	NA	15	UBL7-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	23995761	47	CAGE	ASO G0247240_AD_05	HDF (neonata)
DRR176853	CNhi10785	GCT.experiment	Targeted	NA	15	DNAJC27-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14270899	47	CAGE	ASO G0224165_AD_07	HDF (neonata)
DRR176852	CNhi10785	CAC.experiment	Targeted	NA	16	ACO09005.2	Homo sapiens	FANTOM6 project	PRJDB7993	15040228	47	CAGE	ASO G0267751_AD_02	HDF (neonata)
DRR176787	CNhi10763	AGT.experiment	Targeted	NA	14	RP11-54A9.1	Homo sapiens	FANTOM6 project	PRJDB7993	16807513	47	CAGE	ASO G0257219_AD_10	HDF (neonata)
DRR176786	CNhi10763	ACG.experiment	Negative control	NA	13	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12903573	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176785	CNhi10763	ACC.experiment	Experiment control	NA	14	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	18573976	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR176784	CNhi10762	TAC.experiment	Negative control	NA	13	NA	Homo sapiens	FANTOM6 project	PRJDB7993	22431766	47	CAGE	ASO_NC_B	HDF (neonata)
DRR176783	CNhi10762	GCT.experiment	Negative control	NA	14	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15133454	47	CAGE	ASO_Lipo	HDF (neonata)
DRR176782	CNhi10762	GCG.experiment	Targeted	NA	14	RP11-173B14.5	Homo sapiens	FANTOM6 project	PRJDB7993	15666869	47	CAGE	ASO G0261105_AD_10	HDF (neonata)
DRR176781	CNhi10762	CAC.experiment	Targeted	NA	14	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	10965846	47	CAGE	ASO G0270141_AD_10	HDF (neonata)
DRR176780	CNhi10762	ATG.experiment	Targeted	NA	14	CTD-265I820.3	Homo sapiens	FANTOM6 project	PRJDB7993	23608421	47	CAGE	ASO G0259520_AD_08	HDF (neonata)
DRR176715	CNhi10638	ACC.experiment	Targeted	NA	12	RP6-99M1.2	Homo sapiens	FANTOM6 project	PRJDB7993	15226663	47	CAGE	ASO G0270069_AD_07	HDF (neonata)
DRR176714	CNhi10637	TAC.experiment	Targeted	NA	12	AC017048.4	Homo sapiens	FANTOM6 project	PRJDB7993	18297779	47	CAGE	ASO G0224577_AD_05	HDF (neonata)
DRR176713	CNhi10637	GCT.experiment	Targeted	NA	12	AC017048.3	Homo sapiens	FANTOM6 project	PRJDB7993	14132462	47	CAGE	ASO G0163364_AD_01	HDF (neonata)
DRR176712	CNhi10637	GCG.experiment	Negative control	NA	12	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13491334	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176711	CNhi10637	CAC.experiment	Negative control	NA	12	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12913482	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176710	CNhi10637	ATG.experiment	Targeted	NA	12	TUG1	Homo sapiens	FANTOM6 project	PRJDB7993	13187568	47	CAGE	ASO G0253352_AD_07	HDF (neonata)
DRR176709	CNhi10637	AGT.experiment	Targeted	NA	12	LINC00657	Homo sapiens	FANTOM6 project	PRJDB7993	9691434	47	CAGE	ASO G0260032_AD_01	HDF (neonata)
DRR176708	CNhi10637	ACG.experiment	Targeted	NA	12	ERVX3-1	Homo sapiens	FANTOM6 project	PRJDB7993	17282502	47	CAGE	ASO G0142396_AD_01	HDF (neonata)
DRR176643	CNhi10612	TAC.experiment	Targeted	NA	10	CTA-211A9.5	Homo sapiens	FANTOM6 project	PRJDB7993	12426973	47	CAGE	ASO G0244625_09	HDF (neonata)
DRR176642	CNhi10612	GCT.experiment	Targeted	NA	10	RP11-379H18.1	Homo sapiens	FANTOM6 project	PRJDB7993	11312664	47	CAGE	ASO G0271122_05	HDF (neonata)
DRR176641	CNhi10612	GCG.experiment	Negative control	NA	10	NA	Homo sapiens	FANTOM6 project	PRJDB7993	8554159	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176640	CNhi10612	CAC.experiment	Targeted	NA	10	RP11-110I1.12	Homo sapiens	FANTOM6 project	PRJDB7993	10598063	47	CAGE	ASO G0255121_08	HDF (neonata)
DRR176639	CNhi10612	ATG.experiment	Targeted	NA	10	LINC00707	Homo sapiens	FANTOM6 project	PRJDB7993	10820903	47	CAGE	ASO G0238266_03	HDF (neonata)
DRR176638	CNhi10612	AGT.experiment	Targeted	NA	11	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10342208	47	CAGE	ASO G0246067_AD_09	HDF (neonata)
DRR176637	CNhi10612	ACG.experiment	Targeted	NA	10	CTC-228N24.3	Homo sapiens	FANTOM6 project	PRJDB7993	11000331	47	CAGE	ASO G0245937_01	HDF (neonata)
DRR176636	CNhi10612	ACC.experiment	Targeted	NA	11	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	12006165	47	CAGE	ASO G0246067_AD_06	HDF (neonata)
DRR176571	CNhi10603	GCG.experiment	Targeted	NA	10	SOC52-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	9602476	47	CAGE	ASO G0246985_10	HDF (neonata)
DRR176570	CNhi10603	CAC.experiment	Targeted	NA	10	RP13-463N16.6	Homo sapiens	FANTOM6 project	PRJDB7993	12878134	47	CAGE	ASO G0242147_07	HDF (neonata)
DRR176569	CNhi10603	ATG.experiment	Targeted	NA	10	CKMT2-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10217986	47	CAGE	ASO G0247572_08	HDF (neonata)
DRR176568	CNhi10603	AGT.experiment	Targeted	NA	11	TP53T1	Homo sapiens	FANTOM6 project	PRJDB7993	11620789	47	CAGE	ASO G0182165_AD_08	HDF (neonata)
DRR176567	CNhi10603	ACG.experiment	Targeted	NA	10	SBF2-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10807510	47	CAGE	ASO G0246273_02	HDF (neonata)
DRR176566	CNhi10603	ACC.experiment	Targeted	NA	10	SRP14-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14016720	47	CAGE	ASO G0248508_10	HDF (neonata)
DRR176565	CNhi10602	TAC.experiment	Negative control	NA	11	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15352421	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176564	CNhi10602	GCT.experiment	Targeted	NA	10	RP6-109B7.3	Homo sapiens	FANTOM6 project	PRJDB7993	15350985	47	CAGE	ASO G0241990_07	HDF (neonata)
DRR176499	CNhi10582	ATG.experiment	Targeted	NA	9	RP11-173B14.5	Homo sapiens	FANTOM6 project	PRJDB7993	14226731	47	CAGE	ASO G0261105_02	HDF (neonata)
DRR176498	CNhi10582	AGT.experiment	Targeted	NA	9	RP11-22P6.3	Homo sapiens	FANTOM6 project	PRJDB7993	12327209	47	CAGE	ASO G0260442_01	HDF (neonata)
DRR176497	CNhi10582	ACG.experiment	Targeted	NA	9	RP11-863P13.3	Homo sapiens	FANTOM6 project	PRJDB7993	12325726	47	CAGE	ASO G0261327_05	HDF (neonata)
DRR176496	CNhi10582	ACC.experiment	Targeted	NA	9	RP11-253M7.1	Homo sapiens	FANTOM6 project	PRJDB7993	17087797	47	CAGE	ASO G0259426_02	HDF (neonata)
DRR176495	CNhi10581	TAC.experiment	Targeted	NA	9	FOXN3-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14675723	47	CAGE	ASO G0258920_01	HDF (neonata)
DRR176494	CNhi10581	GCT.experiment	Targeted	NA	9	RP11-253M7.1	Homo sapiens	FANTOM6 project	PRJDB7993	14053872	47	CAGE	ASO G0259426_02	HDF (neonata)

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176493	CNhi10581	GCG.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12106544	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176492	CNhi10581	ACG.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13966469	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176427	CNhi10541	ACG.experiment	Targeted	NA	6	RP11-2201.1	Homo sapiens	FANTOM6 project	PRJDB7993	11123665	47	CAGE	ASO_G0233137_05	HDF (neonata)
DRR176426	CNhi10541	ACC.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11800868	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176425	CNhi10540	TAC.experiment	Targeted	NA	6	MAPKAPK5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11799170	47	CAGE	ASO_G0234608_06	HDF (neonata)
DRR176424	CNhi10540	GCT.experiment	Targeted	NA	5	CLDN22	Homo sapiens	FANTOM6 project	PRJDB7993	13984626	47	CAGE	ASO_C0008202_02	HDF (neonata)
DRR176423	CNhi10540	GCG.experiment	Targeted	NA	5	CATG00000016989.1	Homo sapiens	FANTOM6 project	PRJDB7993	11104412	47	CAGE	ASO_C0020443_03	HDF (neonata)
DRR176422	CNhi10540	CAC.experiment	Targeted	NA	5	RP11-622K12.1	Homo sapiens	FANTOM6 project	PRJDB7993	10844147	47	CAGE	ASO_G0204272_04	HDF (neonata)
DRR176421	CNhi10540	ATG.experiment	Targeted	NA	5	CLDN22	Homo sapiens	FANTOM6 project	PRJDB7993	12571318	47	CAGE	ASO_C0008202_01	HDF (neonata)
DRR176420	CNhi10540	AGT.experiment	Targeted	NA	5	KTNI-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13238245	47	CAGE	ASO_G0186615_09	HDF (neonata)
DRR176355	CNhi10532	ACC.experiment	Targeted	NA	4	RP11-395B7.4	Homo sapiens	FANTOM6 project	PRJDB7993	13374543	47	CAGE	ASO_G0227053_04	HDF (neonata)
DRR176354	CNhi10531	TAC.experiment	Targeted	NA	4	SUGT1P	Homo sapiens	FANTOM6 project	PRJDB7993	12015756	47	CAGE	ASO_G0226823_03	HDF (neonata)
DRR176353	CNhi10531	GCT.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11820256	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176352	CNhi10531	GCG.experiment	Targeted	NA	5	WDFY3-AS2	Homo sapiens	FANTOM6 project	PRJDB7993	8397642	47	CAGE	ASO_G0180769_08	HDF (neonata)
DRR176351	CNhi10531	CAC.experiment	Targeted	NA	5	CATG000000020700.1	Homo sapiens	FANTOM6 project	PRJDB7993	11601366	47	CAGE	ASO_C0020984_02	HDF (neonata)
DRR176350	CNhi10531	ATG.experiment	Targeted	NA	4	RP4-773A18.4	Homo sapiens	FANTOM6 project	PRJDB7993	10808666	47	CAGE	ASO_G0227811_06	HDF (neonata)
DRR176349	CNhi10531	AGT.experiment	Targeted	NA	4	MF12-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11821692	47	CAGE	ASO_G0228109_07	HDF (neonata)
DRR176348	CNhi10531	ACG.experiment	Targeted	NA	4	RP11-395B7.4	Homo sapiens	FANTOM6 project	PRJDB7993	11530324	47	CAGE	ASO_G0227053_03	HDF (neonata)
DRR176283	CNhi10522	GCT.experiment	Targeted	NA	4	PCAT6	Homo sapiens	FANTOM6 project	PRJDB7993	16919426	47	CAGE	ASO_G0228288_01	HDF (neonata)
DRR176282	CNhi10522	GCG.experiment	Negative control	NA	5	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13824230	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176281	CNhi10522	CAC.experiment	Targeted	NA	4	JPX	Homo sapiens	FANTOM6 project	PRJDB7993	16348095	47	CAGE	ASO_G0225470_05	HDF (neonata)
DRR176280	CNhi10522	ATG.experiment	Targeted	NA	4	CTD-2006C1.2	Homo sapiens	FANTOM6 project	PRJDB7993	6267614	47	CAGE	ASO_G0219665_08	HDF (neonata)
DRR176279	CNhi10522	AGT.experiment	Targeted	NA	4	CD99P1	Homo sapiens	FANTOM6 project	PRJDB7993	18248582	47	CAGE	ASO_G0223773_05	HDF (neonata)
DRR176278	CNhi10522	ACG.experiment	Targeted	NA	4	CD27-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	15522154	47	CAGE	ASO_G0215039_01	HDF (neonata)
DRR176277	CNhi10522	ACC.experiment	Targeted	NA	5	AHR	Homo sapiens	FANTOM6 project	PRJDB7993	19880566	47	CAGE	ASO_C0011632_03	HDF (neonata)
DRR176276	CNhi10521	TAC.experiment	Targeted	NA	4	CD27-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	18369099	47	CAGE	ASO_G0215039_01	HDF (neonata)
DRR176211	CNhi10490	GCG.experiment	Targeted	NA	3	AC108488.3	Homo sapiens	FANTOM6 project	PRJDB7993	14460822	47	CAGE	ASO_G0234171_06	HDF (neonata)
DRR176210	CNhi10490	CAC.experiment	Targeted	NA	3	LINC00886	Homo sapiens	FANTOM6 project	PRJDB7993	11307055	47	CAGE	ASO_G0240875_02	HDF (neonata)
DRR176209	CNhi10490	ATG.experiment	Targeted	NA	3	LINC00886	Homo sapiens	FANTOM6 project	PRJDB7993	16000638	47	CAGE	ASO_G0240875_03	HDF (neonata)
DRR176208	CNhi10490	AGT.experiment	Targeted	NA	3	RP11-10K16.1	Homo sapiens	FANTOM6 project	PRJDB7993	17275626	47	CAGE	ASO_G0245213_04	HDF (neonata)
DRR176207	CNhi10490	ACG.experiment	Targeted	NA	3	LINC00856	Homo sapiens	FANTOM6 project	PRJDB7993	13652351	47	CAGE	ASO_G0230417_01	HDF (neonata)
DRR176206	CNhi10490	CAC.experiment	Targeted	NA	3	LINC00517	Homo sapiens	FANTOM6 project	PRJDB7993	14077891	47	CAGE	ASO_G0259091_01	HDF (neonata)
DRR176205	CNhi10489	TAC.experiment	Targeted	NA	3	RP11-500C11.3	Homo sapiens	FANTOM6 project	PRJDB7993	17142828	47	CAGE	ASO_G0272269_06	HDF (neonata)
DRR176204	CNhi10489	GCT.experiment	Negative control	NA	3	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15054983	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176139	CNhi10473	ATG.experiment	Targeted	NA	2	RP11-611E13.2	Homo sapiens	FANTOM6 project	PRJDB7993	19341728	47	CAGE	ASO_G0257815_01	HDF (neonata)
DRR176138	CNhi10473	AGT.experiment	Targeted	NA	2	AC016747.3	Homo sapiens	FANTOM6 project	PRJDB7993	16812682	47	CAGE	ASO_G0212978_02	HDF (neonata)
DRR176137	CNhi10473	ACG.experiment	Targeted	NA	2	AC016747.3	Homo sapiens	FANTOM6 project	PRJDB7993	15343797	47	CAGE	ASO_G0212978_04	HDF (neonata)
DRR176136	CNhi10473	ACC.experiment	Negative control	NA	2	NA	Homo sapiens	FANTOM6 project	PRJDB7993	19771257	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176135	CNhi10472	TAC.experiment	Targeted	NA	2	RP11-398K22.12	Homo sapiens	FANTOM6 project	PRJDB7993	23158002	47	CAGE	ASO_G0229852_05	HDF (neonata)
DRR176134	CNhi10472	GCT.experiment	Targeted	NA	2	RP11-473M20.14	Homo sapiens	FANTOM6 project	PRJDB7993	17414736	47	CAGE	ASO_G0263072_01	HDF (neonata)
DRR176133	CNhi10472	GCG.experiment	Targeted	NA	2	AC016747.3	Homo sapiens	FANTOM6 project	PRJDB7993	15146085	47	CAGE	ASO_G0212978_01	HDF (neonata)
DRR176132	CNhi10472	CAC.experiment	Targeted	NA	2	RP11-473M20.14	Homo sapiens	FANTOM6 project	PRJDB7993	15159810	47	CAGE	ASO_G0263072_05	HDF (neonata)
DRR176067	CNhi10462	ACG.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14049470	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176066	CNhi10462	ACC.experiment	Targeted	NA	1	RP6-99M1.2	Homo sapiens	FANTOM6 project	PRJDB7993	20499075	47	CAGE	ASO_G0270069_05	HDF (neonata)
DRR176065	CNhi10461	TAC.experiment	Targeted	NA	1	RP3-510D11.2	Homo sapiens	FANTOM6 project	PRJDB7993	17807326	47	CAGE	ASO_G0234546_01	HDF (neonata)
DRR176064	CNhi10461	GCT.experiment	Targeted	NA	1	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	17579848	47	CAGE	ASO_G0226950_04	HDF (neonata)
DRR176063	CNhi10461	CAC.experiment	Targeted	NA	1	EMX2OS	Homo sapiens	FANTOM6 project	PRJDB7993	13483938	47	CAGE	ASO_G0229847_03	HDF (neonata)
DRR176062	CNhi10461	ATG.experiment	Targeted	NA	1	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	15651015	47	CAGE	ASO_G0226950_02	HDF (neonata)
DRR176061	CNhi10461	AGT.experiment	Targeted	NA	1	RAB30-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	21575839	47	CAGE	ASO_G0246067_01	HDF (neonata)
DRR176060	CNhi10461	ACG.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13530144	47	CAGE	ASO_NC_A	HDF (neonata)
DRR177139	CNhi10843	TAC.experiment	Negative control	NA	22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	23443023	47	CAGE	sgRNA_NC_gRNA1	HDF (neonata)
DRR177138	CNhi10843	GCT.experiment	Targeted	NA	22	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	12180047	47	CAGE	sgRNA_TERC_gRNA1	HDF (neonata)
DRR177137	CNhi10843	GCG.experiment	Targeted	NA	22	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	12824896	47	CAGE	sgRNA_FTX_gRNA3	HDF (neonata)
DRR176448	CNhi10543	TAC.experiment	Targeted	NA	5	CLDN22	Homo sapiens	FANTOM6 project	PRJDB7993	14503377	47	CAGE	ASO_C0008202_01	HDF (neonata)
DRR176447	CNhi10543	GCT.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14323258	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176446	CNhi10543	GCG.experiment	Targeted	NA	6	MAPKAPK5-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11940335	47	CAGE	ASO_G0234608_06	HDF (neonata)
DRR176445	CNhi10543	CAC.experiment	Targeted	NA	6	ZNF37BP	Homo sapiens	FANTOM6 project	PRJDB7993	11394708	47	CAGE	ASO_G0234420_01	HDF (neonata)
DRR176444	CNhi10543	ATG.experiment	Targeted	NA	5	CATG00000090305.1	Homo sapiens	FANTOM6 project	PRJDB7993	12174880	47	CAGE	ASO_C0011140_06	HDF (neonata)
DRR176379	CNhi10535	ACC.experiment	Targeted	NA	6	AC092295.7	Homo sapiens	FANTOM6 project	PRJDB7993	14698908	47	CAGE	ASO_G0233527_03	HDF (neonata)
DRR176378	CNhi10534	TAC.experiment	Targeted	NA	5	AMZ2P1	Homo sapiens	FANTOM6 project	PRJDB7993	12328135	47	CAGE	ASO_G0214174_06	HDF (neonata)
DRR176377	CNhi10534	GCT.experiment	Experiment control	NA	5	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	15226344	47	CAGE	ASO_MALAT1	HDF (neonata)
DRR176376	CNhi10534	GCG.experiment	Targeted	NA	6	AC012146.7	Homo sapiens	FANTOM6 project	PRJDB7993	8922371	47	CAGE	ASO_G0234327_07	HDF (neonata)
DRR176375	CNhi10534	CAC.experiment	Targeted	NA	5	AMZ2P1	Homo sapiens	FANTOM6 project	PRJDB7993	10882945	47	CAGE	ASO_G0214174_06	HDF (neonata)
DRR176374	CNhi10534	ATG.experiment	Targeted	NA	5	KTNI-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	8214294	47	CAGE	ASO_G0186615_04	HDF (neonata)
DRR176373	CNhi10534	AGT.experiment	Targeted	NA	5	FLVCR1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	12599527	47	CAGE	ASO_G0198468_01	HDF (neonata)
DRR176372	CNhi10534	ACG.experiment	Targeted	NA	5	KTNI-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	11265897	47	CAGE	ASO_G0186615_03	HDF (neonata)
DRR176307	CNhi10525	TAC.experiment	Targeted	NA	4	RP11-527D7.1	Homo sapiens	FANTOM6 project	PRJDB7993	17486887	47	CAGE	ASO_G0225554_02	HDF (neonata)
DRR176306	CNhi10525	GCT.experiment	Targeted	NA	5	RP11-474P2.4	Homo sapiens	FANTOM6 project	PRJDB7993	17799195	47	CAGE	ASO_C0019316_01	HDF (neonata)
DRR176305	CNhi10525	GCG.experiment	Targeted	NA	4	LL22NC03-86G7.1	Homo sapiens	FANTOM6 project	PRJDB7993	14176621	47	CAGE	ASO_G0224086_06	HDF (neonata)
DRR176304	CNhi10525	CAC.experiment	Targeted	NA	5	RP11-474P2.4	Homo sapiens	FANTOM6 project	PRJDB7993	16902977	47	CAGE	ASO_C0019316_08	HDF (neonata)
DRR176303	CNhi10525	ATG.experiment	Targeted	NA	5	RP11-474P2.4	Homo sapiens	FANTOM6 project	PRJDB7993	15755654	47	CAGE	ASO_C0019316_08	HDF (neonata)
DRR176302	CNhi10525	AGT.experiment	Negative control	NA	4	NA	Homo sapiens	FANTOM6 project	PRJDB7993	21533454	47	CAGE	ASO_NC_A	HDF (neonata)
DRR176301	CNhi10525	CAC.experiment	Targeted	NA	4	CD99P1	Homo sapiens	FANTOM6 project	PRJDB7993	15081118	47	CAGE	ASO_G0223773_06	HDF (neonata)
DRR176300	CNhi10525	ACG.experiment	Targeted	NA	5	RP11-474P2.4	Homo sapiens	FANTOM6 project	PRJDB7993	17074292	47	CAGE	ASO_C0019316_01	HDF (neonata)
DRR176235	CNhi10493	GCG.experiment	Targeted	NA	3	AC004980.7	Homo sapiens	FANTOM6 project	PRJDB7993	12570418	47	CAGE	ASO_G0205485_01	HDF (neonata)
DRR176234	CNhi10493	CAC.experiment	Targeted	NA	3	AC108488.3	Homo sapiens	FANTOM6 project	PRJDB7993	10427972	47	CAGE	ASO_G0234171_09	HDF (neonata)
DRR176233	CNhi10493	ATG.experiment	Targeted	NA	3	AC007246.3	Homo sapiens	FANTOM6 project	PRJDB7993	16394140	47	CAGE	ASO_G0231312_02	HDF (neonata)
DRR176232	CNhi10493	AGT.experiment	Targeted	NA	3	ZNF674-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14956403	47	CAGE	ASO_G0230844_01	HDF (neonata)
DRR176231	CNhi10493	ACG.experiment	Targeted	NA	3	AC091729.9	Homo sapiens	FANTOM6 project	PRJDB7993	12494530	47	CAGE	ASO_G0229043_07	HDF (neonata)
DRR176230	CNhi10493	ACC.experiment	Targeted	NA	3	LINC00886	Homo sapiens	FANTOM6 project	PRJDB7993	12134286	47	CAGE	ASO_G0240875_03	HDF (neonata)
DRR176229	CNhi10492	TAC.experiment	Targeted	NA	3	AP001258.4	Homo sapiens	FANTOM6 project	PRJDB7993	19456432	47	CAGE	ASO_G0245571_06	HDF (neonata)
DRR176228	CNhi10492	GCT.experiment	Targeted	NA	3	CTD-2587H24.5	Homo sapiens	FANTOM6 project	PRJDB7993	1591284				

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176163	CNhi10476	CAC.experiment	Targeted	NA	2	RP11-195F19.9	Homo sapiens	FANTOM6 project	PRJDB7993	14704089	47	CAGE	ASO G0230074_07	HDF (neonata)
DRR176162	CNhi10476	ATG.experiment	Targeted	NA	2	RP11-65J3.1	Homo sapiens	FANTOM6 project	PRJDB7993	16087204	47	CAGE	ASO G0233901_04	HDF (neonata)
DRR176161	CNhi10476	ATG.experiment	Targeted	NA	2	BX32257.10	Homo sapiens	FANTOM6 project	PRJDB7993	18493737	47	CAGE	ASO G0215447_01	HDF (neonata)
DRR176160	CNhi10476	ACG.experiment	Targeted	NA	2	A1BG-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14631444	47	CAGE	ASO G0268895_03	HDF (neonata)
DRR176159	CNhi10476	ACG.experiment	Targeted	NA	2	RP11-195F19.9	Homo sapiens	FANTOM6 project	PRJDB7993	20736809	47	CAGE	ASO G0230074_07	HDF (neonata)
DRR176158	CNhi10475	TAC.experiment	Targeted	NA	2	RP11-65J3.1	Homo sapiens	FANTOM6 project	PRJDB7993	21935610	47	CAGE	ASO G0233901_02	HDF (neonata)
DRR176157	CNhi10475	GCT.experiment	Targeted	NA	2	RP11-398K22.12	Homo sapiens	FANTOM6 project	PRJDB7993	20260394	47	CAGE	ASO G0229852_03	HDF (neonata)
DRR176156	CNhi10475	CAC.experiment	Experiment control	NA	2	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	15105188	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176091	CNhi10467	AGT.experiment	Targeted	NA	2	RP11-195F19.9	Homo sapiens	FANTOM6 project	PRJDB7993	18530343	47	CAGE	ASO G0230074_05	HDF (neonata)
DRR176090	CNhi10467	ACG.experiment	Targeted	NA	2	SERTAD4-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	18647594	47	CAGE	ASO G0203706_01	HDF (neonata)
DRR176089	CNhi10467	ACG.experiment	Targeted	NA	2	RP11-458D21.1	Homo sapiens	FANTOM6 project	PRJDB7993	17286834	47	CAGE	ASO G0233396_07	HDF (neonata)
DRR176088	CNhi10466	TAC.experiment	Targeted	NA	2	RP11-150O12.1	Homo sapiens	FANTOM6 project	PRJDB7993	24153957	47	CAGE	ASO C013368_01	HDF (neonata)
DRR176087	CNhi10466	GCT.experiment	Targeted	NA	2	AC027612.6	Homo sapiens	FANTOM6 project	PRJDB7993	16366089	47	CAGE	ASO G0143429_05	HDF (neonata)
DRR176086	CNhi10466	GCG.experiment	Targeted	NA	2	RP11-65J3.1	Homo sapiens	FANTOM6 project	PRJDB7993	13427008	47	CAGE	ASO G0233901_05	HDF (neonata)
DRR176085	CNhi10466	CAC.experiment	Targeted	NA	2	A1BG-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	18970912	47	CAGE	ASO G0268895_03	HDF (neonata)
DRR176084	CNhi10466	ATG.experiment	Targeted	NA	2	RP11-458D21.1	Homo sapiens	FANTOM6 project	PRJDB7993	18474667	47	CAGE	ASO G0233396_07	HDF (neonata)
DRR177091	CNhi10833	ACG.experiment	Negative control	NA	18	NA	Homo sapiens	FANTOM6 project	PRJDB7993	19134187	47	CAGE	ASO NC_A	HDF (neonata)
DRR177090	CNhi10832	TAC.experiment	Targeted	NA	18	CATG0000009599.1	Homo sapiens	FANTOM6 project	PRJDB7993	19516719	47	CAGE	ASO C0012015_AD_02	HDF (neonata)
DRR177089	CNhi10832	GCT.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11832532	47	CAGE	ASO Lipo	HDF (neonata)
DRR177088	CNhi10832	GCG.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	13300348	47	CAGE	ASO NC_A	HDF (neonata)
DRR177087	CNhi10832	CAC.experiment	Targeted	NA	18	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	10384372	47	CAGE	ASO G0226950_04	HDF (neonata)
DRR177086	CNhi10832	ATG.experiment	Targeted	NA	19	CTD-2336O2.1	Homo sapiens	FANTOM6 project	PRJDB7993	27721514	47	CAGE	ASO G0253982_AD_01	HDF (neonata)
DRR177085	CNhi10832	AGT.experiment	Targeted	NA	19	NR2F1-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13530817	47	CAGE	ASO G0237187_AD_01	HDF (neonata)
DRR177084	CNhi10832	ACG.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	17993361	47	CAGE	ASO NC_B	HDF (neonata)
DRR177019	CNhi10815	TAC.experiment	Targeted	NA	20	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	18537536	47	CAGE	siRNA_n269549	HDF (neonata)
DRR177018	CNhi10815	GCT.experiment	Targeted	NA	21	LINC00707	Homo sapiens	FANTOM6 project	PRJDB7993	14406830	47	CAGE	ASO G0238266_AD_06	HDF (neonata)
DRR177017	CNhi10815	GCG.experiment	Targeted	NA	13	RP11-149I23.3	Homo sapiens	FANTOM6 project	PRJDB7993	13854783	47	CAGE	ASO G0232110_AD_06	HDF (neonata)
DRR177016	CNhi10815	CAC.experiment	Targeted	NA	21	HOTAIR	Homo sapiens	FANTOM6 project	PRJDB7993	13379732	47	CAGE	siRNA_ADAAYU4	HDF (neonata)
DRR177015	CNhi10815	ATG.experiment	Negative control	NA	13	NA	Homo sapiens	FANTOM6 project	PRJDB7993	17496519	47	CAGE	ASO NC_A	HDF (neonata)
DRR177014	CNhi10815	AGT.experiment	Targeted	NA	20	EMC7	Homo sapiens	FANTOM6 project	PRJDB7993	17684190	47	CAGE	ASO G0134153_POS_07	HDF (neonata)
DRR177013	CNhi10815	ACG.experiment	Targeted	NA	20	EMC7	Homo sapiens	FANTOM6 project	PRJDB7993	20573720	47	CAGE	ASO G0134153_POS_05	HDF (neonata)
DRR177012	CNhi10815	ACC.experiment	Targeted	NA	17	FOXN3-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	15320993	47	CAGE	ASO G0258920_AD_07	HDF (neonata)
DRR176011	CNhi10454	TAC.experiment	Targeted	NA	1	EMX2OS	Homo sapiens	FANTOM6 project	PRJDB7993	23299491	47	CAGE	ASO G0229847_01	HDF (neonata)
DRR176010	CNhi10454	GCT.experiment	Targeted	NA	1	PTCHD3P1	Homo sapiens	FANTOM6 project	PRJDB7993	17931318	47	CAGE	ASO G0224597_05	HDF (neonata)
DRR176009	CNhi10454	GCG.experiment	Targeted	NA	1	RP11-417E7.1	Homo sapiens	FANTOM6 project	PRJDB7993	14015331	47	CAGE	ASO G0223485_01	HDF (neonata)
DRR176008	CNhi10454	CAC.experiment	Negative control	NA	1	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12960823	47	CAGE	ASO NC_A	HDF (neonata)
DRR176007	CNhi10454	ATG.experiment	Experiment control	NA	1	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	16324581	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176006	CNhi10454	AGT.experiment	Targeted	NA	1	MEG3	Homo sapiens	FANTOM6 project	PRJDB7993	18085898	47	CAGE	ASO G0214548_01	HDF (neonata)
DRR176005	CNhi10454	ACG.experiment	Targeted	NA	1	PTCHD3P1	Homo sapiens	FANTOM6 project	PRJDB7993	14512452	47	CAGE	ASO G0224597_05	HDF (neonata)
DRR176004	CNhi10454	ACG.experiment	Targeted	NA	1	AC017048.4	Homo sapiens	FANTOM6 project	PRJDB7993	17330792	47	CAGE	ASO G0224577_01	HDF (neonata)
DRR176947	CNhi10805	GCT.experiment	Targeted	NA	19	CTD-2336O2.1	Homo sapiens	FANTOM6 project	PRJDB7993	12336277	47	CAGE	ASO G0253982_AD_01	HDF (neonata)
DRR176946	CNhi10805	GCG.experiment	Targeted	NA	19	ZSCAN16-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10376190	47	CAGE	ASO G0269293_AD_10	HDF (neonata)
DRR176945	CNhi10805	CAC.experiment	Targeted	NA	19	RP11-539L10.3	Homo sapiens	FANTOM6 project	PRJDB7993	10497605	47	CAGE	ASO G0251580_AD_05	HDF (neonata)
DRR176944	CNhi10805	ATG.experiment	Targeted	NA	21	AC091729.9	Homo sapiens	FANTOM6 project	PRJDB7993	17459941	47	CAGE	ASO G0229043_AD_04	HDF (neonata)
DRR176943	CNhi10805	AGT.experiment	Targeted	NA	18	AHR	Homo sapiens	FANTOM6 project	PRJDB7993	13044707	47	CAGE	ASO C0011632_AD_01	HDF (neonata)
DRR176942	CNhi10805	ACG.experiment	Targeted	NA	19	RP11-33B1.1	Homo sapiens	FANTOM6 project	PRJDB7993	16466257	47	CAGE	ASO G0245958_AD_07	HDF (neonata)
DRR176941	CNhi10805	ACC.experiment	Negative control	NA	18	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16524020	47	CAGE	ASO Lipo	HDF (neonata)
DRR176940	CNhi10804	TAC.experiment	Negative control	NA	19	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14869760	47	CAGE	ASO NC_A	HDF (neonata)
DRR176875	CNhi10788	CAC.experiment	Targeted	NA	16	LINC00667	Homo sapiens	FANTOM6 project	PRJDB7993	13995528	47	CAGE	ASO G0263753_AD_01	HDF (neonata)
DRR176874	CNhi10788	ATG.experiment	Experiment control	NA	16	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	17985879	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176873	CNhi10788	AGT.experiment	Targeted	NA	16	RP11-333E1.1	Homo sapiens	FANTOM6 project	PRJDB7993	15933723	47	CAGE	ASO G0261879_AD_07	HDF (neonata)
DRR176872	CNhi10788	ACG.experiment	Targeted	NA	17	AC013394.2	Homo sapiens	FANTOM6 project	PRJDB7993	19355784	47	CAGE	ASO G0272888_AD_07	HDF (neonata)
DRR176871	CNhi10788	ACC.experiment	Targeted	NA	15	CTA-211A9.5	Homo sapiens	FANTOM6 project	PRJDB7993	28305515	47	CAGE	ASO G0244625_AD_08	HDF (neonata)
DRR176870	CNhi10787	TAC.experiment	Targeted	NA	16	RP11-333E1.1	Homo sapiens	FANTOM6 project	PRJDB7993	24399594	47	CAGE	ASO G0261879_AD_07	HDF (neonata)
DRR176869	CNhi10787	GCT.experiment	Negative control	NA	16	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15790649	47	CAGE	ASO NC_A	HDF (neonata)
DRR176868	CNhi10787	GCG.experiment	Targeted	NA	16	CTC-559E9.6	Homo sapiens	FANTOM6 project	PRJDB7993	14310408	47	CAGE	ASO G0267419_AD_04	HDF (neonata)
DRR176803	CNhi10765	AGT.experiment	Experiment control	NA	14	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	18093119	47	CAGE	ASO MALAT1	HDF (neonata)
DRR176802	CNhi10765	ACG.experiment	Targeted	NA	14	RP11-553K8.5	Homo sapiens	FANTOM6 project	PRJDB7993	14779060	47	CAGE	ASO G0261573_AD_06	HDF (neonata)
DRR177136	CNhi10843	CAC.experiment	Targeted	NA	22	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	13639309	47	CAGE	sgRNA_TERC_gRNA1_sg	HDF (neonata)
DRR177135	CNhi10843	ATG.experiment	Targeted	NA	22	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	18590338	47	CAGE	sgRNA_FTX_gRNA1_sgRNA	HDF (neonata)
DRR177134	CNhi10843	AGT.experiment	Targeted	NA	22	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	13319765	47	CAGE	sgRNA_FTX_gRNA1	HDF (neonata)
DRR177133	CNhi10843	ACG.experiment	Targeted	NA	22	TERC	Homo sapiens	FANTOM6 project	PRJDB7993	17349932	47	CAGE	sgRNA_TERC_gRNA4	HDF (neonata)
DRR177132	CNhi10842	TAC.experiment	Negative control	NA	22	NA	Homo sapiens	FANTOM6 project	PRJDB7993	15548859	47	CAGE	sgRNA_NC_gRNA1	HDF (neonata)
DRR177067	CNhi10830	ACG.experiment	Targeted	NA	21	RP11-398K22.12	Homo sapiens	FANTOM6 project	PRJDB7993	22569097	47	CAGE	ASO G0229852_AD_01	HDF (neonata)
DRR177066	CNhi10821	TAC.experiment	Negative control	NA	21	NA	Homo sapiens	FANTOM6 project	PRJDB7993	24057977	47	CAGE	siRNA_NC_2	HDF (neonata)
DRR177065	CNhi10821	GCT.experiment	Targeted	NA	20	EXOSC3	Homo sapiens	FANTOM6 project	PRJDB7993	9258485	47	CAGE	siRNA_s27231	HDF (neonata)
DRR177064	CNhi10821	GCG.experiment	Targeted	NA	20	SNRPD3	Homo sapiens	FANTOM6 project	PRJDB7993	10112718	47	CAGE	siRNA_s13235	HDF (neonata)
DRR177063	CNhi10821	CAC.experiment	Targeted	NA	20	EXOSC3	Homo sapiens	FANTOM6 project	PRJDB7993	12765518	47	CAGE	siRNA_s27230	HDF (neonata)
DRR177062	CNhi10821	ACG.experiment	Targeted	NA	20	SNRPD3	Homo sapiens	FANTOM6 project	PRJDB7993	22916540	47	CAGE	siRNA_s13234	HDF (neonata)
DRR177061	CNhi10821	AGT.experiment	Targeted	NA	20	EXOSC10	Homo sapiens	FANTOM6 project	PRJDB7993	10964453	47	CAGE	siRNA_s10738	HDF (neonata)
DRR177060	CNhi10821	ACG.experiment	Targeted	NA	21	CHAF1A	Homo sapiens	FANTOM6 project	PRJDB7993	25145503	47	CAGE	siRNA_s19500	HDF (neonata)
DRR176995	CNhi10812	TAC.experiment	Targeted	NA	14	SOC52-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	26149167	47	CAGE	ASO G0246985_AD_01	HDF (neonata)
DRR176994	CNhi10812	GCT.experiment	Targeted	NA	14	SOC52-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10923390	47	CAGE	ASO G0246985_AD_01	HDF (neonata)
DRR176993	CNhi10812	GCG.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	886737	47	CAGE	ASO Untreat	HDF (neonata)
DRR176992	CNhi10812	CAC.experiment	Targeted	NA	21	RP11-527D7.1	Homo sapiens	FANTOM6 project	PRJDB7993	13969522	47	CAGE	ASO G0225554_AD_06	HDF (neonata)
DRR176991	CNhi10812	ATG.experiment	Targeted	NA	20	CDKN2B-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	16491942	47	CAGE	ASO G0240498_POS_01	HDF (neonata)
DRR176990	CNhi10812	AGT.experiment	Targeted	NA	20	EMC7	Homo sapiens	FANTOM6 project	PRJDB7993	11314569	47	CAGE	siRNA_s32294	HDF (neonata)
DRR176989	CNhi10812	ACG.experiment	Negative control	NA	20	NA	Homo sapiens	FANTOM6 project	PRJDB7993	16628634	47	CAGE	siRNA_NC_1	HDF (neonata)
DRR176988	CNhi10812	ACC.experiment	Targeted	NA	21	HOTAIR	Homo sapiens	FANTOM6 project	PRJDB7993	17575188	47	CAGE	siRNA_AD70LI3	HDF (neonata)

Table_S3 Meatadata for single cell analysis and FANTOM6 kmer research

DRR176923	CNhi10801_GCG.experiment	Experiment control	NA	21	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	10451398	47	CAGE	ASO_MALAT1	HDF (neonata
DRR176922	CNhi10801_CAC.experiment	Negative control	NA	18	NA	Homo sapiens	FANTOM6 project	PRJDB7993	10331467	47	CAGE	ASO_NC_B	HDF (neonata
DRR176921	CNhi10801_ATG.experiment	Targeted	NA	18	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	15780598	47	CAGE	ASO_G0226950_04	HDF (neonata
DRR176920	CNhi10801_AGT.experiment	Targeted	NA	19	NNT-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	14525791	47	CAGE	ASO_G0248092_AD_10	HDF (neonata
DRR176919	CNhi10801_CAC.experiment	Targeted	NA	18	AHR	Homo sapiens	FANTOM6 project	PRJDB7993	17332366	47	CAGE	ASO_C0011632_AD_03	HDF (neonata
DRR176918	CNhi10801_ACC.experiment	Targeted	NA	18	CATG00000095991.1	Homo sapiens	FANTOM6 project	PRJDB7993	16265024	47	CAGE	ASO_C0012015_AD_04	HDF (neonata
DRR176917	CNhi10793_TAC.experiment	Targeted	NA	15	AC009948.5	Homo sapiens	FANTOM6 project	PRJDB7993	17642958	47	CAGE	ASO_G0223960_AD_10	HDF (neonata
DRR176916	CNhi10793_GCT.experiment	Targeted	NA	15	RP5-1103G7.4	Homo sapiens	FANTOM6 project	PRJDB7993	16716427	47	CAGE	ASO_G0226377_AD_05	HDF (neonata
DRR176851	CNhi10785_ATG.experiment	Targeted	NA	15	CTA-211A9.5	Homo sapiens	FANTOM6 project	PRJDB7993	23481414	47	CAGE	ASO_G0244625_AD_08	HDF (neonata
DRR176850	CNhi10785_AGT.experiment	Targeted	NA	16	RP13-20L14.6	Homo sapiens	FANTOM6 project	PRJDB7993	17464058	47	CAGE	ASO_G0265458_AD_03	HDF (neonata
DRR176849	CNhi10785_ACC.experiment	Negative control	NA	16	NA	Homo sapiens	FANTOM6 project	PRJDB7993	23569211	47	CAGE	ASO_Untreat	HDF (neonata
DRR176848	CNhi10785_ACC.experiment	Targeted	NA	17	CTD-3018O17.3	Homo sapiens	FANTOM6 project	PRJDB7993	19110270	47	CAGE	ASO_G0269834_AD_02	HDF (neonata
DRR176847	CNhi10784_TAC.experiment	Targeted	NA	15	CTD-2006C1.2	Homo sapiens	FANTOM6 project	PRJDB7993	20959309	47	CAGE	ASO_G0219665_AD_08	HDF (neonata
DRR176846	CNhi10784_GCT.experiment	Targeted	NA	15	PITPNA-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13142421	47	CAGE	ASO_G0236618_AD_09	HDF (neonata
DRR176845	CNhi10784_GCG.experiment	Negative control	NA	15	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14008049	47	CAGE	ASO_Untreat	HDF (neonata
DRR176844	CNhi10784_CAC.experiment	Targeted	NA	15	UBL7-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	12253203	47	CAGE	ASO_G0247240_AD_06	HDF (neonata
DRR176779	CNhi10762_AGT.experiment	Experiment control	NA	17	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	16038454	47	CAGE	ASO_MALAT1	HDF (neonata
DRR176778	CNhi10762_ACG.experiment	Targeted	NA	14	RP5-886K2.3	Homo sapiens	FANTOM6 project	PRJDB7993	17733373	47	CAGE	ASO_G0236810_AD_06	HDF (neonata
DRR176777	CNhi10762_ACC.experiment	Targeted	NA	13	RP11-295G20.2	Homo sapiens	FANTOM6 project	PRJDB7993	15374841	47	CAGE	ASO_G0233461_AD_03	HDF (neonata
DRR176776	CNhi10761_CAC.experiment	Experiment control	NA	13	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	21544025	47	CAGE	ASO_MALAT1	HDF (neonata
DRR176775	CNhi10761_GCT.experiment	Targeted	NA	13	RP11-149I23.3	Homo sapiens	FANTOM6 project	PRJDB7993	12309090	47	CAGE	ASO_G0232110_AD_02	HDF (neonata
DRR176774	CNhi10761_GCG.experiment	Targeted	NA	13	DANCR	Homo sapiens	FANTOM6 project	PRJDB7993	12748501	47	CAGE	ASO_G0226950_04	HDF (neonata
DRR176773	CNhi10761_CAC.experiment	Negative control	NA	14	NA	Homo sapiens	FANTOM6 project	PRJDB7993	12581492	47	CAGE	ASO_NC_B	HDF (neonata
DRR176772	CNhi10761_ATG.experiment	Targeted	NA	14	RAD51-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	22472899	47	CAGE	ASO_G0245849_AD_07	HDF (neonata
DRR176707	CNhi10637_ACC.experiment	Targeted	NA	12	DNM30S	Homo sapiens	FANTOM6 project	PRJDB7993	15053770	47	CAGE	ASO_G0230630_AD_04	HDF (neonata
DRR176706	CNhi10636_TAC.experiment	Targeted	NA	12	FTX	Homo sapiens	FANTOM6 project	PRJDB7993	16918381	47	CAGE	ASO_G0230590_AD_02	HDF (neonata
DRR176705	CNhi10636_GCT.experiment	Targeted	NA	12	AC017048.3	Homo sapiens	FANTOM6 project	PRJDB7993	16509918	47	CAGE	ASO_G0163364_AD_01	HDF (neonata
DRR176704	CNhi10636_GCG.experiment	Targeted	NA	12	LINC00657	Homo sapiens	FANTOM6 project	PRJDB7993	13636151	47	CAGE	ASO_G0260032_AD_06	HDF (neonata
DRR176703	CNhi10636_CAC.experiment	Targeted	NA	12	LINC00657	Homo sapiens	FANTOM6 project	PRJDB7993	13217433	47	CAGE	ASO_G0260032_AD_01	HDF (neonata
DRR176702	CNhi10636_ATG.experiment	Targeted	NA	12	EMX2OS	Homo sapiens	FANTOM6 project	PRJDB7993	16244806	47	CAGE	ASO_G0229847_AD_02	HDF (neonata
DRR176701	CNhi10636_AGT.experiment	Targeted	NA	12	EMX2OS	Homo sapiens	FANTOM6 project	PRJDB7993	11824543	47	CAGE	ASO_G0229847_AD_02	HDF (neonata
DRR176700	CNhi10636_ACG.experiment	Experiment control	NA	12	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	14694197	47	CAGE	ASO_MALAT1	HDF (neonata
DRR176635	CNhi10611_TAC.experiment	Experiment control	NA	10	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	11501566	47	CAGE	ASO_MALAT1	HDF (neonata
DRR176634	CNhi10611_GCT.experiment	Targeted	NA	10	RP11-110I1.12	Homo sapiens	FANTOM6 project	PRJDB7993	8406979	47	CAGE	ASO_G0255121_08	HDF (neonata
DRR176633	CNhi10611_GCG.experiment	Negative control	NA	10	NA	Homo sapiens	FANTOM6 project	PRJDB7993	9274218	47	CAGE	ASO_NC_A	HDF (neonata
DRR176632	CNhi10611_CAC.experiment	Targeted	NA	10	RP6-109B7.3	Homo sapiens	FANTOM6 project	PRJDB7993	10469369	47	CAGE	ASO_G0241990_01	HDF (neonata
DRR176631	CNhi10611_ATG.experiment	Targeted	NA	10	AC013394.2	Homo sapiens	FANTOM6 project	PRJDB7993	9125788	47	CAGE	ASO_G0272888_04	HDF (neonata
DRR176630	CNhi10611_AGT.experiment	Targeted	NA	10	CTA-211A9.5	Homo sapiens	FANTOM6 project	PRJDB7993	12678086	47	CAGE	ASO_G0244625_09	HDF (neonata
DRR176629	CNhi10611_ACG.experiment	Targeted	NA	10	CTA-211A9.5	Homo sapiens	FANTOM6 project	PRJDB7993	11573367	47	CAGE	ASO_G0244625_04	HDF (neonata
DRR176628	CNhi10611_ACC.experiment	Targeted	NA	10	SOCS2-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10980987	47	CAGE	ASO_G0246985_10	HDF (neonata
DRR176563	CNhi10602_GCG.experiment	Negative control	NA	10	NA	Homo sapiens	FANTOM6 project	PRJDB7993	10065128	47	CAGE	ASO_NC_A	HDF (neonata
DRR176562	CNhi10602_CAC.experiment	Negative control	NA	10	NA	Homo sapiens	FANTOM6 project	PRJDB7993	10476544	47	CAGE	ASO_NC_A	HDF (neonata
DRR176561	CNhi10602_ATG.experiment	Targeted	NA	10	CTD-3018O17.3	Homo sapiens	FANTOM6 project	PRJDB7993	12001193	47	CAGE	ASO_G0269834_02	HDF (neonata
DRR176560	CNhi10602_AGT.experiment	Negative control	NA	10	NA	Homo sapiens	FANTOM6 project	PRJDB7993	10821049	47	CAGE	ASO_NC_A	HDF (neonata
DRR176559	CNhi10602_ACG.experiment	Targeted	NA	10	RP11-379H18.1	Homo sapiens	FANTOM6 project	PRJDB7993	12325227	47	CAGE	ASO_G0271122_06	HDF (neonata
DRR176558	CNhi10602_ACC.experiment	Targeted	NA	10	SRP14-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	12510693	47	CAGE	ASO_G0248508_04	HDF (neonata
DRR176557	CNhi10601_TAC.experiment	Targeted	NA	11	LINC00883	Homo sapiens	FANTOM6 project	PRJDB7993	17599940	47	CAGE	ASO_G0243701_AD_08	HDF (neonata
DRR176556	CNhi10601_GCT.experiment	Targeted	NA	10	SRP14-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	13194337	47	CAGE	ASO_G0248508_04	HDF (neonata
DRR176491	CNhi10581_ATG.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	14020684	47	CAGE	ASO_NC_A	HDF (neonata
DRR176490	CNhi10581_AGT.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	11403096	47	CAGE	ASO_NC_A	HDF (neonata
DRR176489	CNhi10581_ACG.experiment	Targeted	NA	9	FOXN3-AS1	Homo sapiens	FANTOM6 project	PRJDB7993	10725925	47	CAGE	ASO_G0258920_01	HDF (neonata
DRR176488	CNhi10581_ACC.experiment	Experiment control	NA	6	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	12558472	47	CAGE	ASO_MALAT1	HDF (neonata
DRR176487	CNhi10580_TAC.experiment	Targeted	NA	6	RP11-149I23.3	Homo sapiens	FANTOM6 project	PRJDB7993	10467445	47	CAGE	ASO_G0232110_04	HDF (neonata
DRR176486	CNhi10580_GCT.experiment	Targeted	NA	6	RP11-295G20.2	Homo sapiens	FANTOM6 project	PRJDB7993	10687172	47	CAGE	ASO_G0233461_03	HDF (neonata
DRR176485	CNhi10580_GCG.experiment	Targeted	NA	6	RP11-149I23.3	Homo sapiens	FANTOM6 project	PRJDB7993	9900277	47	CAGE	ASO_G0232110_04	HDF (neonata
DRR176484	CNhi10580_CAC.experiment	Targeted	NA	6	RP4-794H19.4	Homo sapiens	FANTOM6 project	PRJDB7993	9965639	47	CAGE	ASO_G0230812_10	HDF (neonata
DRR176419	CNhi10540_ACG.experiment	Negative control	NA	6	NA	Homo sapiens	FANTOM6 project	PRJDB7993	10262820	47	CAGE	ASO_NC_A	HDF (neonata
DRR176418	CNhi10540_ACC.experiment	Targeted	NA	5	CATG00000016989.1	Homo sapiens	FANTOM6 project	PRJDB7993	11327335	47	CAGE	ASO_C0020443_06	HDF (neonata
DRR176417	CNhi10539_TAC.experiment	Targeted	NA	5	CLDN22	Homo sapiens	FANTOM6 project	PRJDB7993	17879932	47	CAGE	ASO_C0008202_02	HDF (neonata
DRR176416	CNhi10539_GCT.experiment	Targeted	NA	6	AC092295.7	Homo sapiens	FANTOM6 project	PRJDB7993	15872598	47	CAGE	ASO_G0233527_02	HDF (neonata
DRR176415	CNhi10539_GCG.experiment	Targeted	NA	5	CATG00000090305.1	Homo sapiens	FANTOM6 project	PRJDB7993	10645579	47	CAGE	ASO_C0011140_02	HDF (neonata
DRR176414	CNhi10539_CAC.experiment	Targeted	NA	5	CATG00000090305.1	Homo sapiens	FANTOM6 project	PRJDB7993	13275292	47	CAGE	ASO_C0011140_06	HDF (neonata
DRR176413	CNhi10539_ATG.experiment	Targeted	NA	6	MAGI2-AS3	Homo sapiens	FANTOM6 project	PRJDB7993	13121092	47	CAGE	ASO_G0234456_01	HDF (neonata
DRR176412	CNhi10539_AGT.experiment	Experiment control	NA	5	MALAT1	Homo sapiens	FANTOM6 project	PRJDB7993	11547941	47	CAGE	ASO_MALAT1	HDF (neonata
DRR176347	CNhi10531_ACC.experiment	Targeted	NA	4	RP1-152L7.5	Homo sapiens	FANTOM6 project	PRJDB7993	11790821	47	CAGE	ASO_G0216775_01	HDF (neonata
DRR176346	CNhi10530_TAC.experiment	Targeted	NA	5	AHR	Homo sapiens	FANTOM6 project	PRJDB7993	10482319	47	CAGE	ASO_C0011632_10	HDF (neonata

Table_S4 metadata for kmer research in MSC in different conditions

Run ID	Run Title	Experiment ID	EXTERNAL_ID	Specie	study_title	project ID	nreads	read_length	cell type	passages	treatment
SRR5878104	GSM2720343: ATF6+/- hMSC vehicle-2 RNA-seq; Homo sapiens; RNA-Seq	SRX3045407	GSM2720343	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	2991402	101	MSC	p5	vehicle
SRR5878103	GSM2720342: ATF6+/- hMSC vehicle-1 RNA-seq; Homo sapiens; RNA-Seq	SRX3045406	GSM2720342	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	2957936	101	MSC	p5	vehicle
SRR5878102	GSM2720341: ATF6+/- hMSC tunicamycin-2 RNA-seq; Homo sapiens; RNA-Seq	SRX3045405	GSM2720341	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	11944303	101	MSC	p5	tunicamycin
SRR5878101	GSM2720340: ATF6+/- hMSC tunicamycin-1 RNA-seq; Homo sapiens; RNA-Seq	SRX3045404	GSM2720340	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	9700764	101	MSC	p5	tunicamycin
SRR5878100	GSM2720339: ATF6-/- hMSC tunicamycin-2 RNA-seq; Homo sapiens; RNA-Seq	SRX3045403	GSM2720339	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	11592138	101	MSC	p5	tunicamycin
SRR5878099	GSM2720338: ATF6-/- hMSC tunicamycin-1 RNA-seq; Homo sapiens; RNA-Seq	SRX3045402	GSM2720338	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	10540607	101	MSC	p5	tunicamycin
SRR5878098	GSM2720337: ATF6-/- hMSC vehicle-2 RNA-seq; Homo sapiens; RNA-Seq	SRX3045401	GSM2720337	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	9985873	101	MSC	p5	vehicle
SRR5878097	GSM2720336: ATF6-/- hMSC vehicle-1 RNA-seq; Homo sapiens; RNA-Seq	SRX3045400	GSM2720336	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	9924834	101	MSC	p5	vehicle
SRR5878096	GSM2720335: ATF6+/- hWAPC vehicle-2 RNA-seq; Homo sapiens; RNA-Seq	SRX3045399	GSM2720335	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	32613469	101	WAPC	p5	vehicle
SRR5878095	GSM2720334: ATF6+/- hWAPC vehicle-1 RNA-seq; Homo sapiens; RNA-Seq	SRX3045398	GSM2720334	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	11973184	101	WAPC	p5	vehicle
SRR5878094	GSM2720333: ATF6+/- hWAPC tunicamycin-2 RNA-seq; Homo sapiens; RNA-Seq	SRX3045397	GSM2720333	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	10129610	101	WAPC	p5	tunicamycin
SRR5878093	GSM2720332: ATF6+/- hWAPC tunicamycin-1 RNA-seq; Homo sapiens; RNA-Seq	SRX3045396	GSM2720332	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	12628610	101	WAPC	p5	tunicamycin
SRR5878092	GSM2720331: ATF6-/- hWAPC tunicamycin-2 RNA-seq; Homo sapiens; RNA-Seq	SRX3045395	GSM2720331	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	10730944	101	WAPC	p5	tunicamycin
SRR5878091	GSM2720330: ATF6-/- hWAPC tunicamycin-1 RNA-seq; Homo sapiens; RNA-Seq	SRX3045394	GSM2720330	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	9530313	101	WAPC	p5	tunicamycin
SRR5878090	GSM2720329: ATF6-/- hWAPC vehicle-2 RNA-seq; Homo sapiens; RNA-Seq	SRX3045393	GSM2720329	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	10217050	101	WAPC	p5	vehicle
SRR5878089	GSM2720328: ATF6-/- hWAPC vehicle-1 RNA-seq; Homo sapiens; RNA-Seq	SRX3045392	GSM2720328	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	10306775	101	WAPC	p5	vehicle
SRR5878080	GSM2720319: ATF6+/- hMSC late passage-2 RNA-seq; Homo sapiens; RNA-Seq	SRX3045383	GSM2720319	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	8182392	143.61	MSC	p10	NA
SRR5878079	GSM2720318: ATF6+/- hMSC late passage-1 RNA-seq; Homo sapiens; RNA-Seq	SRX3045382	GSM2720318	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	6753082	147.91	MSC	p10	NA
SRR5878078	GSM2720317: ATF6-/- hMSC late passage-2 RNA-seq; Homo sapiens; RNA-Seq	SRX3045381	GSM2720317	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	7782173	147.47	MSC	p10	NA
SRR5878077	GSM2720316: ATF6-/- hMSC late passage-1 RNA-seq; Homo sapiens; RNA-Seq	SRX3045380	GSM2720316	Homo sapiens	ATF6 safeguards organelle homeostasis and cellular aging in human mesenchymal stem cells	PRJNA396193	8331869	145.24	MSC	p10	NA
Run ID	Run Title	Experiment ID	EXTERNAL_ID	Specie	study_title	project ID	nreads	read_length	cell type	passages	treatment
SRR6680317	GSM2983811: hMSC_TAZ_3; Homo sapiens; RNA-Seq	SRX3656633	GSM2983811	Homo sapiens	Upregulation of FOXD1 by YAP alleviates senescence and osteoarthritis	PRJNA433339	10429618	150	MSC		TAZ KO
SRR6680316	GSM2983810: hMSC_TAZ_2; Homo sapiens; RNA-Seq	SRX3656632	GSM2983810	Homo sapiens	Upregulation of FOXD1 by YAP alleviates senescence and osteoarthritis	PRJNA433339	11433785	150	MSC		TAZ KO
SRR6680315	GSM2983809: hMSC_TAZ_1; Homo sapiens; RNA-Seq	SRX3656631	GSM2983809	Homo sapiens	Upregulation of FOXD1 by YAP alleviates senescence and osteoarthritis	PRJNA433339	11009653	150	MSC		TAZ KO
SRR6680314	GSM2983808: hMSC_YAP_3; Homo sapiens; RNA-Seq	SRX3656630	GSM2983808	Homo sapiens	Upregulation of FOXD1 by YAP alleviates senescence and osteoarthritis	PRJNA433339	11196860	150	MSC		YAP KO
SRR6680313	GSM2983807: hMSC_YAP_2; Homo sapiens; RNA-Seq	SRX3656629	GSM2983807	Homo sapiens	Upregulation of FOXD1 by YAP alleviates senescence and osteoarthritis	PRJNA433339	11261647	150	MSC		YAP KO
SRR6680312	GSM2983806: hMSC_YAP_1; Homo sapiens; RNA-Seq	SRX3656628	GSM2983806	Homo sapiens	Upregulation of FOXD1 by YAP alleviates senescence and osteoarthritis	PRJNA433339	10926717	150	MSC		YAP KO
SRR6680311	GSM2983805: hMSC_WT_3; Homo sapiens; RNA-Seq	SRX3656627	GSM2983805	Homo sapiens	Upregulation of FOXD1 by YAP alleviates senescence and osteoarthritis	PRJNA433339	11048601	150	MSC		WT
SRR6680310	GSM2983804: hMSC_WT_2; Homo sapiens; RNA-Seq	SRX3656626	GSM2983804	Homo sapiens	Upregulation of FOXD1 by YAP alleviates senescence and osteoarthritis	PRJNA433339	10156337	150	MSC		WT
SRR6680309	GSM2983803: hMSC_WT_1; Homo sapiens; RNA-Seq	SRX3656625	GSM2983803	Homo sapiens	Upregulation of FOXD1 by YAP alleviates senescence and osteoarthritis	PRJNA433339	10341635	150	MSC		WT
Run ID	Run Title	Experiment ID	EXTERNAL_ID	Specie	study_title	project ID	nreads	read_length	cell type	passages	treatment
SRR5357818	GSM2544142: MSC 211 CytoD; Homo sapiens; RNA-Seq	SRX2653659	GSM2544142	Homo sapiens	Osteogenic programming of adipose-derived mesenchymal stem cells using a fungal metabolite that suppresses the Polycomb protein EZH2	PRJNA379707	25484965	51	adipose derived mesenchymal stem cells	5-7	CytoD
SRR5357817	GSM2544141: MSC 211 Ctrl; Homo sapiens; RNA-Seq	SRX2653658	GSM2544141	Homo sapiens	Osteogenic programming of adipose-derived mesenchymal stem cells using a fungal metabolite that suppresses the Polycomb protein EZH2	PRJNA379707	26464456	51	adipose derived mesenchymal stem cells	5-7	Ctrl

Table_S4 metadata for kmer research in MSC in different conditions

SRR5357816	GSM2544140: MSC 283 CytoD; Homo sapiens; RNA-Seq	SRX2653657	GSM2544140	Homo sapiens	Osteogenic programming of adipose-derived mesenchymal stem cells using a fungal metabolite that suppresses the Polycomb protein EZH2	PRJNA379707	24280737	51	adipose derived mesenchymal stem cells	5-7	CytoD
SRR5357815	GSM2544139: MSC 283 Ctrl; Homo sapiens; RNA-Seq	SRX2653656	GSM2544139	Homo sapiens	Osteogenic programming of adipose-derived mesenchymal stem cells using a fungal metabolite that suppresses the Polycomb protein EZH2	PRJNA379707	21040712	51	adipose derived mesenchymal stem cells	5-7	Ctrl
SRR5357814	GSM2544138: MSC 258 CytoD; Homo sapiens; RNA-Seq	SRX2653655	GSM2544138	Homo sapiens	Osteogenic programming of adipose-derived mesenchymal stem cells using a fungal metabolite that suppresses the Polycomb protein EZH2	PRJNA379707	35139902	51	adipose derived mesenchymal stem cells	5-7	CytoD
SRR5357813	GSM2544137: MSC 258 Ctrl; Homo sapiens; RNA-Seq	SRX2653654	GSM2544137	Homo sapiens	Osteogenic programming of adipose-derived mesenchymal stem cells using a fungal metabolite that suppresses the Polycomb protein EZH2	PRJNA379707	34623656	51	adipose derived mesenchymal stem cells	5-7	Ctrl
Run ID	Run Title	Experiment ID	EXTERNAL_ID	specie	study_title	project ID	nreads	read_length	cell type	passages	treatment
SRR8447838	GSM3564546: 10 days diff Control siRNA-C RNA-seq; Homo sapiens; RNA-Seq	SRX5254873	SAMN10755603	Homo sapiens	PLZF targets developmental enhancers for activation during osteogenic differentiation of human mesenchymal stem cells (RNA-seq)	PRJNA515466	17969174	44	human Mesenchymal Stem Cells	p4-p10	10d osteo Diff
SRR8447837	GSM3564545: 10 days diff Control siRNA-B RNA-seq; Homo sapiens; RNA-Seq	SRX5254872	SAMN10755604	Homo sapiens	PLZF targets developmental enhancers for activation during osteogenic differentiation of human mesenchymal stem cells (RNA-seq)	PRJNA515466	18943085	44	human Mesenchymal Stem Cells	p4-p10	10d osteo Diff
SRR8447836	GSM3564544: 10 days diff Control siRNA-A RNA-seq; Homo sapiens; RNA-Seq	SRX5254871	SAMN10755605	Homo sapiens	PLZF targets developmental enhancers for activation during osteogenic differentiation of human mesenchymal stem cells (RNA-seq)	PRJNA515466	26461685	44	human Mesenchymal Stem Cells	p4-p10	10d osteo Diff
SRR8447835	GSM3564543: 2 days diff Control siRNA-C RNA-seq; Homo sapiens; RNA-Seq	SRX5254870	SAMN10755606	Homo sapiens	PLZF targets developmental enhancers for activation during osteogenic differentiation of human mesenchymal stem cells (RNA-seq)	PRJNA515466	27117845	44	human Mesenchymal Stem Cells	p4-p10	2d osteo Diff
SRR8447834	GSM3564542: 2 days diff Control siRNA-B RNA-seq; Homo sapiens; RNA-Seq	SRX5254869	SAMN10755607	Homo sapiens	PLZF targets developmental enhancers for activation during osteogenic differentiation of human mesenchymal stem cells (RNA-seq)	PRJNA515466	26162294	44	human Mesenchymal Stem Cells	p4-p10	2d osteo Diff
SRR8447833	GSM3564541: 2 days diff Control siRNA-A RNA-seq; Homo sapiens; RNA-Seq	SRX5254868	SAMN10755608	Homo sapiens	PLZF targets developmental enhancers for activation during osteogenic differentiation of human mesenchymal stem cells (RNA-seq)	PRJNA515466	19724867	44	human Mesenchymal Stem Cells	p4-p10	2d osteo Diff
SRR8447832	GSM3564540: NaiveControl siRNA-C RNA-seq; Homo sapiens; RNA-Seq	SRX5254867	SAMN10755609	Homo sapiens	PLZF targets developmental enhancers for activation during osteogenic differentiation of human mesenchymal stem cells (RNA-seq)	PRJNA515466	24266576	44	human Mesenchymal Stem Cells	p4-p10	Naive
SRR8447831	GSM3564539: NaiveControl siRNA-B RNA-seq; Homo sapiens; RNA-Seq	SRX5254866	SAMN10755610	Homo sapiens	PLZF targets developmental enhancers for activation during osteogenic differentiation of human mesenchymal stem cells (RNA-seq)	PRJNA515466	25910627	44	human Mesenchymal Stem Cells	p4-p10	Naive
SRR8447830	GSM3564538: NaiveControl siRNA-A RNA-seq; Homo sapiens; RNA-Seq	SRX5254865	SAMN10755611	Homo sapiens	PLZF targets developmental enhancers for activation during osteogenic differentiation of human mesenchymal stem cells (RNA-seq)	PRJNA515466	24849706	44	human Mesenchymal Stem Cells	p4-p10	Naive
Run ID	Run Title	Experiment ID	EXTERNAL_ID	specie	study_title	project ID	nreads	read_length	cell type	passages	treatment
SRR3882937	GSM2231443: AMSC_19_proliferating; Homo sapiens; RNA-Seq	SRX1939121	GSM2231443	Homo sapiens	Transcriptome Sequencing of Adipose-Derived Mesenchymal Stromal Cells	PRJNA328824	29447967	50	Adipose-derived Mesenchymal Stromal Cell		
SRR3882936	GSM2231442: AMSC_19_confluent; Homo sapiens; RNA-Seq	SRX1939120	GSM2231442	Homo sapiens	Transcriptome Sequencing of Adipose-Derived Mesenchymal Stromal Cells	PRJNA328824	43754165	51	Adipose-derived Mesenchymal Stromal Cell		
SRR3882935	GSM2231441: AMSC_4_proliferating; Homo sapiens; RNA-Seq	SRX1939119	GSM2231441	Homo sapiens	Transcriptome Sequencing of Adipose-Derived Mesenchymal Stromal Cells	PRJNA328824	34428851	50	Adipose-derived Mesenchymal Stromal Cell		
SRR3882934	GSM2231440: AMSC_4_confluent; Homo sapiens; RNA-Seq	SRX1939118	GSM2231440	Homo sapiens	Transcriptome Sequencing of Adipose-Derived Mesenchymal Stromal Cells	PRJNA328824	42327600	51	Adipose-derived Mesenchymal Stromal Cell		
SRR3882933	GSM2231439: AMSC_3_proliferating; Homo sapiens; RNA-Seq	SRX1939117	GSM2231439	Homo sapiens	Transcriptome Sequencing of Adipose-Derived Mesenchymal Stromal Cells	PRJNA328824	33767562	51	Adipose-derived Mesenchymal Stromal Cell		
SRR3882932	GSM2231438: AMSC_3_confluent; Homo sapiens; RNA-Seq	SRX1939116	GSM2231438	Homo sapiens	Transcriptome Sequencing of Adipose-Derived Mesenchymal Stromal Cells	PRJNA328824	44801550	51	Adipose-derived Mesenchymal Stromal Cell		
SRR3882931	GSM2231437: AMSC_1_proliferating; Homo sapiens; RNA-Seq	SRX1939115	GSM2231437	Homo sapiens	Transcriptome Sequencing of Adipose-Derived Mesenchymal Stromal Cells	PRJNA328824	37501459	50	Adipose-derived Mesenchymal Stromal Cell		
SRR3882930	GSM2231436: AMSC_1_confluent; Homo sapiens; RNA-Seq	SRX1939114	GSM2231436	Homo sapiens	Transcriptome Sequencing of Adipose-Derived Mesenchymal Stromal Cells	PRJNA328824	34863048	51	Adipose-derived Mesenchymal Stromal Cell		
Run ID	Run Title	Experiment ID	EXTERNAL_ID	specie	study_title	project ID	nreads	read_length	cell type	source_name	treatment
SRR8099005	GSM3444217: EGF_IWR1_S03; Homo sapiens; RNA-Seq	SRX4925800	GSM3444217	Homo sapiens	Epidermal growth factor activates β-catenin via integrin-linked kinase to control proliferation of mesenchymal stromal cells.	PRJNA498109	14005865	151	Mesenchymal stromal cells	femoral heads/knee bone explants	EGF + IWR-1
SRR8099004	GSM3444216: EGF_IWR1_S02; Homo sapiens; RNA-Seq	SRX4925799	GSM3444216	Homo sapiens	Epidermal growth factor activates β-catenin via integrin-linked kinase to control proliferation of mesenchymal stromal cells.	PRJNA498109	11575790	151	Mesenchymal stromal cells	femoral heads/knee bone explants	EGF + IWR-1
SRR8099003	GSM3444215: EGF_IWR1_S01; Homo sapiens; RNA-Seq	SRX4925798	GSM3444215	Homo sapiens	Epidermal growth factor activates β-catenin via integrin-linked kinase to control proliferation of mesenchymal stromal cells.	PRJNA498109	15666602	151	Mesenchymal stromal cells	femoral heads/knee bone explants	EGF + IWR-1
SRR8099002	GSM3444214: EGF_S03; Homo sapiens; RNA-Seq	SRX4925797	GSM3444214	Homo sapiens	Epidermal growth factor activates β-catenin via integrin-linked kinase to control proliferation of mesenchymal stromal cells.	PRJNA498109	15786160	151	Mesenchymal stromal cells	femoral heads/knee bone explants	EGF
SRR8099001	GSM3444213: EGF_S02; Homo sapiens; RNA-Seq	SRX4925796	GSM3444213	Homo sapiens	Epidermal growth factor activates β-catenin via integrin-linked kinase to control proliferation of mesenchymal stromal cells.	PRJNA498109	19843529	151	Mesenchymal stromal cells	femoral heads/knee bone explants	EGF
SRR8099000	GSM3444212: EGF_S01; Homo sapiens; RNA-Seq	SRX4925795	GSM3444212	Homo sapiens	Epidermal growth factor activates β-catenin via integrin-linked kinase to control proliferation of mesenchymal stromal cells.	PRJNA498109	19132337	151	Mesenchymal stromal cells	femoral heads/knee bone explants	EGF
SRR8098999	GSM3444211: untreated_S03; Homo sapiens; RNA-Seq	SRX4925794	GSM3444211	Homo sapiens	Epidermal growth factor activates β-catenin via integrin-linked kinase to control proliferation of mesenchymal stromal cells.	PRJNA498109	13571062	151	Mesenchymal stromal cells	femoral heads/knee bone explants	untreated
SRR8098998	GSM3444210: untreated_S02; Homo sapiens; RNA-Seq	SRX4925793	GSM3444210	Homo sapiens	Epidermal growth factor activates β-catenin via integrin-linked kinase to control proliferation of mesenchymal stromal cells.	PRJNA498109	16677487	151	Mesenchymal stromal cells	femoral heads/knee bone explants	untreated
SRR8098997	GSM3444209: untreated_S01; Homo sapiens; RNA-Seq	SRX4925792	GSM3444209	Homo sapiens	Epidermal growth factor activates β-catenin via integrin-linked kinase to control proliferation of mesenchymal stromal cells.	PRJNA498109	17232973	151	Mesenchymal stromal cells	femoral heads/knee bone explants	untreated

Table_S5 genes names positively enriched in MSC

gene_name
KRTAP1.5
FAM180A
GALNT5
CCKAR
LINC01423
KCNK2
ITGBL1
APCDD1L
BDKRB2
FGF7
LRRCL15
PTGFR
AL355607.2
KCNK15
FGF5
SLC14A1
MYLK
SMILR
GDF5
SBSN
RFX8
CA12
IL13RA2
EYA2
C2CD6
DLX5
PENK
AL359834.1
C1S
RAB27B
Z97200.1
AC022467.1
AL160153.1
FAP
NTM
FAM133CP
FOXD1
C11orf87
ISLR
SFRP4
HOXC10
PRR16
HOXC11
AC091182.2
FAM225B
LINC02454
LAMA3
THBS2
PDGFRA
LIMA1
AC007336.3
BICC1
OLFML1
LINC01605
KIAA1755
AC105046.1
IGFBP6
SATB2
PDGFRB
CLCA2
GPR1
BLID
EGR1
FGF10
PAMR1
LINC00968
PCOLCE
POPC3
C1R
CCN5
COL8A1
SSPN
APCDD1L.DT
AGAP11
SOC55
AL355102.1
AP003071.4
MOXD1
CLMP
GLYATL2
NTNG1
AC114284.1

Table_S5 genes names positively enriched in MSC

LMO7
 RF01891
 PQLC2L
 CCDC68
 TNFRSF11B
 COL6A3
 PTHLH
 PRRX1
 PAM
 PALM2.AKAP2
 SIX2
 AC099066.2
 MIR31HG
 PDE1C
 SLC16A4
 VGLL3
 EGR3
 ABCA8
 SERPINE2
 ITGA11
 IL1RL2
 PTGES
 AC010533.1
 ANXA10
 FOXF2
 HSPB2
 WNT5A
 PPFIBP1
 CCN6
 AL118522.1
 GREM1
 HOTAIR
 FMN2
 BEND6
 ROS1
 IL6
 AC041040.1
 AC110597.1
 RANBP3L
 LMO7.AS1
 LINC01614
 PABPC1P4
 ANTXR2
 SVIL
 DDR2
 KANK2
 HSPB2.C11orf52
 MIR3609
 LTBP2
 COL14A1
 NCAM2
 TSLP
 NEGR1
 HAS1
 ANGPTL5
 DCBLD2
 Z99289.2
 CCN4
 FLG
 IFNE
 ADAMTS6
 XG
 ABCA6
 SEMA5A
 COL6A2
 LINC01060
 SNAI2
 SHISAL1
 SCARA5
 FNDC1
 TNC
 PAPP
 CYGB
 ADAM12
 FAM87B
 TMEM119
 GPR39
 PTGIS
 LEPR
 BX005019.1
 TMEM233
 PCDHGA12
 VIT

Table_S5 genes names positively enriched in MSC

DSEL
DIRAS3
GLT8D2
CFAP300
AC012531.2
PKD2
IGFBP3
MEDAG
PNLIPRP3
ANKRD29
ADAMTS5
PLAT
CXCL12
TFPI2
KCND2
CCDC80
TMEM200A
SVEP1
ADGRL4
MFAP5
CYBRD1
ALDH1A3
RPL35P6
STC1
AC093627.6
IRX3
FIBIN
CHL1
NEAT1
STEAP2
SYNPO2
GFRA1
ETV1
TBX15
PLA2R1
LMOD1
NTSE
PDE7B
HOXC9
ECM2
NEXN
ADAMTS1
LAYN
VSTM4
SIX1
AOX1
ANKRD35
C1GALT1
AC079949.2
SERTAD4
SCUBE3
TWIST1
HSPB7
VEGFC
CD248
STMN2
THY1
LINC02029
AC087821.1
MAP1A
TMEM158
ZNF385D
STEAP1
DYNC2H1
SPOCK1
GEM
MRGPRF
EBF1
LBP
WNT5B
CDCP1
FAM3C
LOXL1
HSPB6
GXYLT2
RNU6.1045P
AC012531.1
GBE1
COL12A1
HOXC6
BAALC
C11orf96
KCTD16

Table_S5 genes names positively enriched in MSC

DAAM2
 CFAP299
 LRRN4CL
 CALHM5
 CDR1
 LINC00601
 LPAR1
 HEG1
 AHI1
 HAS2
 LIF
 CTHRC1
 AC006058.1
 PHLDA1
 FOSL1
 RNF182
 RNU6.530P
 RBPJ
 OLFM2L2A
 EYA1
 AC112721.2
 LYNX1
 ADAMTS2
 DCN
 EDIL3
 AC234772.2
 MAP4K4
 TIMP1
 COL1A2
 CRYAB
 PSD3
 AL359095.1
 TNFAIP6
 CIQTNF1
 CTSK
 GLIS1
 PPL
 GREM2
 NLRP10
 SMIM3
 ITGB1
 PRRX2
 PAPPAS2
 ANKRD28
 FMO2
 HSPD1P11
 GLRB
 MEOX2
 ZCCHC24
 MKX
 NFIX
 AC022034.1
 AC113383.1
 TRAM2
 IRX5
 CCDC110
 PEAK1
 DIO2
 SCG2
 VEGFA
 BMERB1
 C8orf48
 MIR100HG
 CDH11
 MCFD2
 GAS1RR
 AC107308.1
 EPS8
 FOXC2
 CARMN
 SLIT2.IT1
 GALNT1
 USP53
 SPATA18
 ENDOD1
 SNED1
 B3GNT9
 SGIP1
 AC011586.2
 SLC5A3
 COL10A1
 ABCA9
 LINC02202

Table_S5 genes names positively enriched in MSC

PACERR
 PLA2G5
 FBN1
 SOD3
 COL6A1
 RORB
 GABRE
 ABI3BP
 ACTA2
 MIR137HG
 CORO2B
 ADIRF
 TLCD2
 LAMA4
 CALB2
 OLFML3
 INHBA
 ZFHX4
 GDNF
 ARHGAP21
 FRMPD4
 GSC
 BDNF
 NFIC
 HIF1A
 PCDHGA7
 TAGLN
 TGFBI
 AL121748.1
 BDKRB1
 KY
 KCNS2
 AC015922.3
 HTRA1
 CTF1
 Z99289.3
 CFH
 HHIPL2
 NPR3
 LINC01705
 ELOVL3
 OSMR
 KCTD4
 ITGA10
 SGCD
 SLC25A27
 ECM1
 SH2D4A
 EMILIN1
 LRRTM2
 LINC01119
 ID4
 DST
 LURAP1L
 STEAP4
 AC117489.1
 GDF6
 PEXSL
 TIMP3
 ANTXR1
 MIR222HG
 GAS6.DT
 CNTNAP4
 MXRA5
 SERTAD4.AS1
 HOXC5
 SYTL5
 IFITM10
 DOK5
 ARID5B
 SLC1A1
 HEPH
 NID2
 SNORD114.1
 FHL2
 TBX2
 HOXC8
 PDZRN3
 AP000525.1
 CCBE1
 REP15
 SYNC
 MTIL

Table_S5 genes names positively enriched in MSC

HGF
FRMD6
RHOBTB3
PYGO1
TBX18
DGKI
SEC24AP1
NAALADL2
TMEM132B
NA.
AC092807.3
BAG2
GSTM2
VASN
POM121L9P
ZFHX4.AS1
CIART
FAM20A
FGF1
SLC2A12
AC027335.1
EPHA5
RAB23
CEP170
CPEB1
EDNRA
DNAL1
NNMT
AC016397.1
AC007673.1
LUM
CDKN2B
MYL9
IL7
RCAN1
SGCB
PTGS2
KCNJ15
EPDR1
ZMAT3
BNIP3L
VCAN
RN7SKP11
BRINP1
EBF3
LTBP3
QSOX1
PPP1R3C
HMCN1
SCN2A
BMPER
COL16A1
TENM4
AC058822.1
PXK
PLPPR4
POSTN
FOXC1
SNX9
RPSAP52
SATB2.AS1
TPM2
TTC8
LINC02511
AP001434.1
RASSF8
PRKD1
GRAMD2B
LINC01050
SNAP25
AC092691.1
ASPH
PPFIA2
SHOX2
GRIA3
RFTN2
LHFPL6
RNU6.1226P
CASC4
SRPX
FAM114A1
NBL1
TM9SF3

Table_S5 genes names positively enriched in MSC

PCDHGB7
 COMP
 LRP1
 CMYA5
 TWIST2
 MICU1
 NEK7
 LEP
 PALM2
 PALM2
 SNAPC1
 RSUP3
 INMT.MINDY4
 SNHG18
 SPAG16
 SH3PXD2A
 ALPK2
 EPB41L2
 LINC00973
 MEIS3P1
 CLGN
 PCOLCE2
 AC134043.2
 PCDHGA3
 FST
 GAS6
 IGFBP4
 ADAMTSL1
 SERP2
 ARHGAP20
 PROS1
 SYNDIG1
 EMX2OS
 SFRP1
 CFAP69
 UACA
 NPHP3.ACAD11
 LTBP1
 SETD7
 HRH1
 FHOD3
 GAS1
 GGT5
 PCDHGB3
 BHLHE41
 ABCC9
 SUGCT
 MTMR11
 CEMIP
 TFP1
 EEF1B2P3
 KCNJ8
 ANGPT1
 EFEMP2
 RGS4
 CCDC89
 SYAP1
 AC112229.3
 EGFR
 LINC01139
 EGFLAM
 FMOD
 AL359091.2
 LYPD6B
 PCDHGA10
 STEAP1B
 AHR
 SULF1
 TNFSF11
 HOXC4
 SRPX2
 EHD2
 CST1
 MXRA7
 PNMA2
 TREM1
 IGFBP5
 TMEM26
 AP000526.1
 FGF2
 KCNMA1
 SMURF2
 AC093908.1
 KTN1

Table_S5 genes names positively enriched in MSC

CUL4B
 CALD1
 IGF1
 MSRB3
 NTRK2
 GRIN2A
 EYA4
 GALNT15
 FMO3
 MAGI2
 AC096537.1
 DYNC112
 CPXM2
 DTWD1
 TMEM100
 AC112198.2
 MXRA8
 NFIL3
 ATP8B1
 NR4A1
 NME5
 NET1
 AC092645.1
 AGTR1
 AC107021.2
 SSC5D
 CHI3L1
 PCDHB7
 PABPC5
 RPL13AP5
 ARHGAP12
 CCN3
 MMP2
 TNKS1BP1
 STXBP5
 CRNDE
 PLEKHH2
 JCAD
 DDO
 RGS17
 AC013268.1
 PTX3
 CAV1
 ANGPTL2
 NFASC
 SLC6A15
 WDR63
 LINC00632
 SLC4A4
 LNCSRLR
 DLX2
 MXI1
 SMIM10
 PCDH18
 P4HA1
 BTC
 OLFML2B
 NHS
 LINC00839
 CREB3L1
 NAV3
 PCDHGB5
 STXBP1
 HMGB1P21
 DOCK5
 RIMS1
 PRPH2
 ARHGAP6
 PDGFC
 CERCAM
 TWSG1
 MIR8485
 FER1L6
 NKX3.2
 PRUNE2
 HYDIN2
 PCDHGC3
 TMEM47
 ERFF1
 MLPH
 AC008163.1
 FZD1
 PCDHGA2

Table_S5 genes names positively enriched in MSC

HOXA13
 LOXL2
 GJB2
 RASSF9
 NUCB2
 TIPARP
 AL391822.1
 AC026124.1
 DUXAP8
 SNORD114.7
 GFPT2
 CS0138
 CCDC74B
 CPA4
 BHMT2
 AP000892.3
 CEP126
 LRATD1
 CCDC170
 SPTLC3
 ANKH
 TK2
 OSR1
 EGLN1
 GLIS3
 NOTCH3
 C14orf39
 PCDHGA4
 IFI44L
 RHOQ
 RF00181
 EXT1
 TSC22D1
 SMC02
 GNA14
 EFS
 CCPG1
 PCDHGB2
 BVES
 ZC2HC1A
 ARMC9
 ALDH1L2
 AC073591.1
 TCEAL7
 FAT1
 GJA1
 SEMA3C
 FGFR1
 CPEB2
 AC006213.4
 PLCE1.AS1
 IFIT1
 DNM3OS
 ENPP2
 AC007750.1
 AC002480.1
 AZI2
 FAM155A.IT1
 OSR2
 TRIOBP
 CFAP36
 FNDCC3B
 AKR1C2
 Z93022.1
 VDR
 ADORA1
 GPC6
 PTPRM
 AC012321.1
 THAP12P7
 DSTN
 ESM1
 FUT11
 RCAN2
 COL1A1
 REEP3
 KIAA1549L
 AC004264.1
 FSTL1
 EMP1
 C4orf47
 ITPRIPL2
 SNORD114.14

Table_S5 genes names positively enriched in MSC

MYPN
 PLEKHA5
 TENT5A
 BANK1
 PODN
 C8orf88
 IL1R1
 IGFBP7
 ZNF280D
 C12orf56
 ADAMTS9.AS2
 AC080038.1
 AC016397.2
 KCNS1
 NR2F1.AS1
 AEBP1
 AC022509.3
 IRX6
 SRGAP1
 EDA2R
 LINC01106
 PCDHGC5
 FOXL1
 GLS
 LINC02381
 TMEM92.AS1
 CAST
 LINC01117
 NRN1
 HECW2
 RPS12P5
 FN1
 AC073115.2
 PDLIM5
 SCARA3
 TRPC4
 RAET1G
 TMEM171
 NR0B1
 HOXC.AS2
 GPR176
 MRPS6
 ANO3
 FAM171B
 FAM155A
 SPAG9
 MYO1B
 VCAM1
 CYP1B1
 FBXO32
 COL21A1
 AL035078.1
 AL354811.1
 KCNE4
 CHRDL1
 AC009093.1
 SLC9A7
 FAM13C
 CYP2U1
 RNF180
 MME
 FANK1
 EHBP1
 PRICKLE2
 SLC39A6
 CACNA2D1
 THSD4
 SOX9
 TNFRSF19
 PDE5A
 GOLGA2
 DYRK3
 DLC1
 NDUFA4L2
 C14orf132
 EVC
 AC138969.1
 HMG2.AS1
 SLC2A10
 AL137026.1
 CEP290
 SNX33
 SMOC1

Table_S5 genes names positively enriched in MSC

AC093535.2
LXN
ACOX2
BNC2
NALCN
LOXL4
ACSS3
NBPFF10
CAV2
MAP3K20
DIXDC1
RECK
MET
AC013565.1
MAGI2.AS3
FER
VCAN.AS1
TCEAL4
SNORD114.4
CACNA1A
G0S2
PLAC9
RNU6.786P
KLF10
PLSCR4
BCL2L2.PABPN1
COL11A1
TTLL7
HDLBP
TP63
ANGPTL1
ZNF214
TNS2
TDRD6
NEK10
NRP1
CSGALNACT2
TCEAL3
HHIPL1
SPOCD1
ADD3
CHST3
DSE
EGR2
TM4SF1
CACNA1C
ARHGAP22
MIR22HG
AXL
ARSJ
BNC1
B3GALNT1
LARP6
ELL2
CAPN2
PTPN14
DIRAS1
AKR1C1
SEMA3A
HK2
ENAH
COR22
WDR19
CLDN11
RND3
LRP11
NGF
CD109
BX322650.1
VEGFB
TLN2
SEC24D
ARNT2
FKBP7
ATOX8
NFATC4
MEG3
TMEM263
PRCD
PCYOX1
ASAP3
SMAD6
PCDHGA11

Table_S5 genes names positively enriched in MSC

GPX8
 BMS1P7
 HSD11B1
 AL157702.2
 HCG11
 ARHGAP23
 RPS3AP6
 WDFY3.AS2
 AC244669.2
 TPBG
 SLC22A3
 COL3A1
 LRP12
 CDH13
 ROR1
 SLC20A1
 KRT16
 NA.1
 EEA1
 AC096531.2
 KIRREL1
 CDKN2A
 ANK2
 ACAN
 ENOX1
 REXO2
 CEP63
 AC008269.1
 AC233263.6
 RNF24
 LNCTAM34A
 HTR7
 AC090527.2
 VEPH1
 LINC00702
 NRXN2
 PDE1A
 DZIP1
 MAB21L2
 TMTC1
 NKAPL
 CCND1
 LAMC1
 FAM126A
 ITGA7
 SNORD114.2
 ISM1
 CCDC158
 TCEAL9
 LINC01303
 TRH
 CYS1
 AC008522.1
 PARP3
 LDHA
 MAP9
 UHRF2
 ABCB5
 NRG2
 RILPL1
 EREG
 DYNLT3
 PLOD2
 RNASE4
 ABCA13
 AKR1C3
 ANKRD30B
 NEK11
 RNU7.119P
 ARMH4
 PDE3A
 FOXC2.AS1
 IKBIP
 ARHGEF12
 ANKRD53
 SUCLG2.AS1
 RN7SL396P
 MDM2
 IFIT3
 PPP1R14C
 FAS
 AKAP2
 AC069079.1

Table_S5 genes names positively enriched in MSC

CLEC3B
PTCHD4
CSF1
TCAF1
FAM225A
CALU
AL162171.1
B4GALT1
PCDHGA6
DLX6
MTCL1
PPARGC1A
BHLHE40
WWTR1
AC098650.1
SCARB2
C19orf12
TIMP2
CDK15
CCDC8
SPIRE1
TTC3
FKBP9
HFE
STARD13.IT1
P4HA3
AC245297.1
AMOTL1
RAB13
KITLG
CCDC74A
LINC00595
CCN1
PRSS23
ZNF471
HOXA10
AP001528.2
ENTPD3
NTN4
BIVM
CAMK2N1
AC093899.2
PDE4DIP
CCL28
CSPG4
FOSB
FYCO1
AP001970.1
MMP16
NPHP1
DPT
DUXAP10
CAVIN1
HERC4
MIR635
UBE2E3
CRTAP
ADAMTS14
TMCO3
STK17A
ARHGAP24
AL356432.3
TMF1
NPM1P40
IFT46
STC2
TMEM246
TPM1
SNORD113.2
AL157935.2
AC027277.1
TBC1D19
ST13P18
KAZALD1
NPIPA9
PRKG1
ITGB8
FLOT1
ZBTB4
AC140479.4
GSTM5
SERPINF1
PTPRG

Table_S5 genes names positively enriched in MSC

NAP1L1
 AHNK2
 SPRY4
 LCA5
 CHMP3
 PRKAR1A
 ROCK2
 GOLGB1
 HAS2.AS1
 AC090229.1
 TPM4
 PAQR5
 FBXL7
 AC107982.1
 DRP2
 CTSB
 AL049838.1
 ITPRIP
 RAB31
 MINDY2
 TICAM2
 AC037198.1
 SHC1
 NAB2
 RARG
 BARX1
 LOX
 RF00181.1
 BMPR2
 PLXNA4
 KIAA1614
 LDB2
 POGLUT3
 AL391427.1
 SNORD114.6
 NKX3.1
 SGM52
 NR2F2.AS1
 CA9
 SEC22B
 SMIM14
 PCDHGB1
 BCAP29
 AC048341.1
 FZD7
 GNG12
 GAP43
 CH25H
 THBS3
 PPFIA4
 RPS6KA6
 THBS1
 CCL13
 GDNF.AS1
 AC008914.1
 IRS1
 EEF1A1P16
 ARMCX2
 DNAJB4
 EMP2
 STS
 PCOLCE.AS1
 SORBS2
 TRAM1L1
 MSANTD3
 STRIP2
 NDP
 CAMK2D
 MEG8
 SPESP1
 FAM66B
 SLC16A2
 BBX1.AS1
 KLF4
 AC010655.3
 WIP1
 PARVA
 IFT81
 AK5
 CC2D2A
 LINC00472
 GLIS2
 COX20

Table_S5 genes names positively enriched in MSC

PCDHGB4
COLEC12
RAI14
CEP162
NRG1
ITGAV
RPSAP48
SEMA3E
PLCE1
MYOF
RSPH3
LRRC32
NAV1
ARMCX3
TRPA1
AC121247.1
MAP4K5
DUXAP9
SNTB2
RTN4
LDOC1
AL590560.2
LINC02516
SERPINE1
PPIL6
AC015712.2
TMED7.TICAM2
CLIP3
DLX1
CPZ
SETP6
SERINC1
IGIP
RNU6.767P
SEC16B
SLIT3
CPED1
GINM1
FAM20C
FZD8
CTSO
YAP1
PRG4
ESR1
PTN
CCN2
CYP27C1
RASD2
TGFB2
TRIO
ZNF229
STAT2
TMTC3
CFL2
DCUN1D3
TTC30A
PDCD1LG2
PPP3CA
ARHGAP35
EFEMP1
GOLIM4
ACSM5
AC079594.2
AC025423.1
CHSY3
MOK
AL021068.1
DZIP3
PLPP7
KIAA1217
TGFB1I1
MIR1245A
FAM210B
HMG2P5
TCEA3
AC025259.3
DEPDC7
LUCAT1
RBFOX2
ERO1A
EMX2
ZHX3
LINC00565

Table_S5 genes names positively enriched in MSC

SLC4A7
ATP2B4
TLR3
LINC01293
PTPN20
MBNL1.AS1
COL25A1
SLC13A4
ABLI M3
STS
HCFC1R1
AL022068.1
MYOZ2
AC008771.1
PTPRQ
MBOAT2
AC099494.2
SHISA4
LSM1P1
AL137024.1
HRCT1
AC100854.1
BACE1
EN1
AC021205.1
RBMS2
CDKL2
SLPI
AC092111.2
EVC2
RCN3
UNC5C
PRELP
C12orf75
AL117190.1
IFI6
TTC3P1
CTTN
FAM129B
AC110597.3
HTR2A
SAMD11
ARMCX1
TEAD1
STAC
PHLDB1
FGF14
SPEG
CBR3
ITGA1
SIX4
HTATSF1P2
SPRY4.AS1
OSBPL8
SEC31A
PTPRK
DLX3
ZNF662
NR4A2
ITGB5
C10orf198
RP1
KIF1BP
MBTPS2
TP1P1
RABGEF1
PRDM16
AL137161.1
PTPRN
DTNA
AFF4
MRV1
MEF2A
AC027031.2
FGFRL1
NCKAP5
GCSHP3
STX1A
ACP7
LRRRC66
ZDHHC7
FOS
RBMS3

Table_S5 genes names positively enriched in MSC

AC108734.4
PER3
CLCF1
NEK1
LINC01123
MT1M
DKK1
LAMB2
AC105114.1
AL049795.2
YWHAE
ZBTB38
RF00100
FIP1L1
AL133346.1
AD000090.1
LIN7C
PPP1R12A
TSHZ3
THSD1
EIF2AK2
TBX2.AS1
PRDM8
SLC41A1
PTGDS
ARHGEF28
MIR497HG
CDH2
AC084757.3
PDLIM4
DKK3
RMDN2
PAPPA.AS1
C1RL
LAMB1
MFSB1
EML1
MTMR9LP
CNTLN
MMP14
ARHGAP22.IT1
GNPNAT1
FZD6
RPLP0P2
AC011611.3
LACC1
UGDH
ZFPM2.AS1
SNTG2.AS1
CDK14
MCC
CYB5R3
APP
AC026785.2
DCLK1
RRM2B
GLIPR1
WNT16
PCSK1
THNSL2
MIEF1
PLPP3
KCTD1
KIF5B
SPIN1
AF127577.4
SGCE
EOGT
RASA1
SNORD113.1
NCOA7
DCDC1
ADH1B
CCDC149
HNRNP1P1
UNC5B
C5orf24
AC110048.2
DAB2
MAP1B
MAMLD1
BOC
PTPN13

Table_S5 genes names positively enriched in MSC

VCL
NFAT5
XYLT1
LINC01844
SYNM
RWDD1
MAT2A
SLC30A4
NR4A3
MFAP4
NDRG1
AC241952.1
RGN
IFIT5
PRICKLE1
ANKRD42
FERMT2
ZFP91
CSPG4P12
RGMB
HOXA10.AS
EXTL2
TULP3
TSPYL5
LINC01268
SNORD114.3
COL4A2
SLIT2
HES1
IFI16
BNC2.AS1
FBLN1
TWFB
NFYB
RAB9B
AKAP11
ENPP1
AL163636.2
RCN1P2
DCUN1D4
EVA1A
LINC01600
PPP2R3A
TUSC3
FOXQ1
A4GALT
ATP9A
FEZ1
FAM167B
C7
SMAD9
LRIG3
RN7SL602P
RCN1
NBPF14
CTTNBP2NL
PRSS3
PPA2
ERMAP
LRRRC17
ALCAM
CHD9
N4BP2L2
LRRFIP2
AL137782.1
ZBTB47
AL162386.2
LINC00882
AL645922.1
KCNT2
CNIH3
AC092104.1
NPY2R
HLX
AMOTL2
PLCD4
SAMD8
NKX6.1
AR
AC104051.2
AC013271.1
FAM187A
AL592158.1

Table_S5 genes names positively enriched in MSC

BGN
 VPS41
 CFAP97
 MALL
 APPL2
 CSNK1G3
 SERPING1
 TCEAL1
 ROBO1
 STEAP3
 LRRC6
 BBS10
 PRSS12
 DUSP4
 FEM1C
 RAET1E
 NDUFS4
 PKP1
 SNX7
 ZBTB41
 RECQL
 GHR
 AL583722.1
 AC090197.1
 CPS1
 MPP4
 ENG
 NPTX1
 KIAA0408
 ZC3HAV1L
 PLOD1
 PLAU
 SDAD1P1
 B3GALT2
 ULBP2
 LAPTM4A
 AC020909.4
 SAP30
 ZNF25
 LAMA1
 SIGLEC15
 AC004080.6
 ATXN3
 JAK2
 IER5L
 SYNJ2
 EBF2
 SLC2A3P1
 RPL23AP32
 PHEX
 PPIL4
 RN7SKP255
 SYT12
 CRISPLD2
 AL078599.2
 MLF1
 HOXD8
 HOXD.AS2
 MARCH4
 ATP10D
 SERINC3
 NDNF
 EXTL1
 SMAD7
 FAT4
 OPCML
 ADM
 ZNF302
 LIFR
 PIFO
 ARL4C
 RF00017
 LGALS3BP
 WDR60
 ENPP5
 AC142381.3
 ARRD3
 FLNC
 MIR4458HG
 PODNL1
 LRRC49
 SPRED2
 NORAD

Table_S5 genes names positively enriched in MSC

AL445649.1
 AC108463.1
 ZNF454
 PHLDB2
 PCDHGA9
 MFAP2
 EPGN
 SMAGP
 SUSD5
 INSYN2A
 STON1
 AC105105.1
 AC023908.1
 KIRREL1.IT1
 AC092614.1
 PXDN
 CCDC14
 PRRT2
 RBM43
 BBOF1
 LEPROT
 CCNG2
 CRYBG3
 TAF7L
 TCTN1
 CAVIN3
 PTPRU
 ADCY8
 HTR1F
 METTL15P1
 WASHC2C
 ADGRL2
 FPGT.TNNI3K
 COL5A2
 DIP2C
 AC092168.2
 ZNF135
 TRIM22
 ITGA3
 RMDN2.AS1
 C11orf58
 PLCB4
 SERPINB2
 TMED2
 AC073115.1
 OSMR.AS1
 SEPT11
 RUNX2
 AC103796.1
 AC080038.2
 LINC01443
 LRP1B
 CCDC36
 GIGYF2
 AP003396.5
 DZIP1L
 KIF7
 AL132780.2
 MPDZ
 PRTFDC1
 ADAM9
 SOX5
 SLC8A1
 LINC02595
 EPAS1
 SNORD113.3
 RARRES1
 PHAX
 ZNF503
 CLIC4
 PHYH
 AL160408.4
 HOXA11
 LINC01426
 PCDHA3
 LMNA
 VAX2
 KIAA0825
 SPAGL7
 TPD52L1
 HOXA7
 H6PD
 KIAA1109

Table_S5 genes names positively enriched in MSC

AC112229.1
CDC14B
RPS6KA2
COL4A4
PRRG1
CYP39A1
SOCS6
DDAH1
MYO1E
TOPORS
FAM110B
LYPD1
MDH1B
AC092053.3
PSG5
ACKR4
FGF7P6
FBLN7
PCDHGA5
MAP7D3
ERLIN2
LRRK2
KCNH1
WDR78
C1GALT1C1L
C15orf65
NCKAP1
MIA3
FUT8
SOBP
TTC37
TCTN2
CHMP2B
MARVELD1
LOXL1.AS1
SH3RF3
ATL3
AC008147.2
WASF3
DPYSL3
ATP2B1
COL24A1
TSPAN11
AC079328.2
DLEU2L
GOLM1
AC009237.3
ELOVL4
LRCH2
RPL23AP7
BX679664.3
ADCY6
CFI
STARD13
PKM
RUNX1T1
FKBP10
SOCS3
ZNF221
GPNMB
ENPEP
TRIB1
LAMA2
HAND2
SYNE1
ADAM33
TM4SF20
AC009237.14
SPX
AC096636.1
RHBDL2
RTL5
RALGPS2
WFDC21P
LINC00327
SLC44A1
SESTD1
SNRNP27
OXR1
NR1D2
AC027237.4
TBX3
ACTRT3

Table_S5 genes names positively enriched in MSC

BACE2
 AC002480.2
 COL5A1
 CRIM1
 SASIH1
 HMGB3P9
 ST13
 GLI3
 AC112721.1
 ITGA2
 UGGT2
 AC068590.4
 TMEM136
 FAM172A
 FKBP14
 CEBPD
 KRCC1
 LINC01133
 TMEM200B
 PGM1
 C3orf67
 TBC1D2
 BMPR1A
 RASSF8.AS1
 MNS1
 HMGA2
 LY96
 KDELR3
 KLF9
 APLN
 TSPAN5
 ZSCAN23
 COL5A3
 PKIG
 SIAE
 STAR13.AS
 SNX25
 AL596325.2
 KIF26B
 CLSTN2
 ELN
 CLU
 FBLN5
 FER1L5
 FJX1
 VKORC1
 SREK1
 BASP1
 ZNF667
 IL20RA
 ZRANB2
 AL590428.1
 RN7SL68P
 PCDH9
 AC040160.1
 ODF2L
 HEXB
 PRDM16.DT
 LIN7A
 GPR161
 MRAS
 CRYZL2P.SEC16B
 RUFY1
 RF00181.2
 NR2F2
 ADCY4
 LRRC27
 PARD3B
 ESF1
 FLG.AS1
 SMTN
 CCDC148
 AL354740.1
 IFI44
 LRRC2
 OSBP1A
 GLI2
 SH3RF1
 PLAUR
 AKAP12
 MYH15
 PCDHGB9P
 PJA2

Table_S5 genes names positively enriched in MSC

AC005840.1
SEMA7A
JUN
SYDE1
FOXP2
FAXDC2
ACTN1
LINC02257
NXPH4
CALCOCO1
CARO6
FAM3C2
CNTNAP1
ABCA1
UBE2H
CNTN3
DYNC2L1
AC005618.1
B3GALT5.AS1
PPIC
PMP22
PALLD
SPRED1
BACH1
LGALS3
AC144831.1
PGRMC2
MT2A
CCDC122
JDP2
DUSP10
AIMP1
ANXA2P2
TMEM255B
AC023043.1
PCDHGB6
TJP1
DHX32
NTSR1
ZDHHC21
MFGE8
TMX3
CDKN2B.AS1
NPR2
AC097534.2
CLDN12
CSDE1
ZFP36L1
TPST1
MSC.AS1
CDKN2C
JCHAIN
ZDHHC1
RPS27L
SEC62
AC079600.3
AC060766.7
TMEFF2
ARHGAP29
SLC25A43
IER3
ZNF436
AC087893.1
ASAP2
RN7SL15P
MMP19
ARNTL2
AL049775.1
DNAJC13
DCAF6
LINC02593
CTSF
CADPS2
CDH6
LINC02016
RPL12P38
EHD3
SLFN5
KIF13A
TMCO1
CPT1C
AOPEP
AC013564.1

Table_S5 genes names positively enriched in MSC

ABL2
 NA..2
 EFNA5
 KIFAP3
 HS3ST3A1
 FRMD5
 FLRT2
 NPC1
 SERTAD2
 SLK
 CTSB
 NA..3
 ZNF300P1
 AC092139.1
 PSORS1C1
 ADGRA2
 IL13RA1
 EFCAB1
 PLA2G12A
 NECTIN3
 ENO1P3
 CCDC92
 IFT80
 P3H2
 DUSP5
 SBDS
 MICU3
 RN7SKP163
 LINC00184
 LGR4
 ZNF106
 HSP90B1
 CD27.AS1
 PCDHGA1
 COX7A1
 PNMA8A
 SLC2A13
 LIMCH1
 HSPA8P4
 EDEM3
 OSS
 ABRAHAS1
 S100A16
 MYBPH
 S100A13
 YAP1P1
 USO1
 OXCT1
 PLK2
 NLGN1
 ADAMTS12
 OTUD7B
 AC046143.1
 AC091965.5
 AC005592.1
 AC073610.2
 TVP23CP1
 WDFY2
 SPART
 CCL26
 FNIP1
 LINC02333
 AP000769.1
 SAR1A
 CNRIP1
 POLR3D
 GTF2H2C
 TXNDC15
 SPRY2
 TEAD3
 CLCN3
 SLC38A2
 SLC35G2
 AP3S1
 CD68
 SLC26A4
 CDKN1A
 AHRR
 REV3L
 CTBS
 ARSK
 LINC02518
 PHC2

Table_S5 genes names positively enriched in MSC

KCNK1
CNN1
AP4E1
FOXP1
BAG3
SLC41A2
GPC1
ANXA1
TSHZ2
BEX3
ID2.AS1
RTL8B
EVI5
PRELID2
AC010198.2
BTBD8
FTH1P23
ENO1P4
RAB3B
SLC25A46
MAGI1
KATNAL2
AC092040.1
MAPK3
TBC1D8B
COPB2
UNC5B.AS1
POLI
ZNF883
AC010201.2
S100A6
RNU4ATAC
AC239809.3
RBMS1
AP000915.2
RTN4RL1
INSY2B
DPYSL2
PCDHB6
STAM2
C2CD2
GPRCSA
RAB36
AP000695.1
RAB10
TEX9
RNF11
TMEM237
AC008758.5
PLEKHA4
PEA15
ZNF281
SMARCA1
NEO1
PXDC1
ATE1.AS1
ARL4AP5
TMEM127
RN7SKP97
ACSL4
ADGRB2
CD59
NEDD8.MDP1
ZNF239
PRNP
IQGAP1
YBX3
C1orf54
CAPS2
PHLDA3
IL11
ZNF626
PGR
NUPR1
DACT1
FTH1P8
EPHA3
APLF
DENND2C
CCDC113
COL13A1
CFAP53
Z97653.2

Table_S5 genes names positively enriched in MSC

AC090204.1
ITGB3
PIGF
FXR1
PLAG1
RPGRIPI1L
SDCCAG8
ETV5
APBB2
AMZ2
CAMLG
P4HA2
KCNIP4.IT1
LINC01275
TRPC1
EEF1A1P7
ETAA1
YIPF5
LANCL3
SSX2IP
RPL5P34
RPL7P1
GNG11
CSDC2
AP001453.2
ZNF354C
AC080038.4
PTEN
NBEAL1
RN7SL718P
AC073508.2
FGD1
MAPK10
SYT11
AC016526.1
TMEM35A
COL8A2
IRX2
IFT57
FRA10AC1
CD01
IER3.AS1
SBF2
KCTD12
CNN3
RASD1
CASC2
INKA2.AS1
TNS1
ATP13A3
RNF152
FBXO17
DPP4
HTR2B
MEIS1
COL4A1
AC074212.1
C7orf25
WASHC2A
PTPRS
ATLP2
GULP1
AL133415.1
WLS
ADH5
LNX1
SPEF2
PLA2G4A
RTL8C
AC124276.2
MIR34AHG
AVPI1
AL133453.1
PCDHB14
TMEM65
F2RL2
TRPS1
FOSL2
ANPEP
RF00181.3
CRISPLD1
AL590438.1
MID1

Table_S5 genes names positively enriched in MSC

SLITRK4
CBLN3
WASHC3
HSPA4L
CCDC82
ZNF415P1
SEPT10
RIN2
LINC02377
MYO10
LINC00571
TUB
SPRED3
PLPP4
GNAI1
TRO
SH3BGR
PCDHGC4
RHOD
MIR604
PRAF2
OAT
RRBP1
AC027117.1
P3H3
PAPSS2
GPER1
METTL7A
AC090617.5
SIL1
ANKIB1
C22orf23
OPTN
CTSL
MSL3P1
AC080013.4
TRIP11
ARHGFE25
ZMYND11
MEGF8
ZSCAN30
AC109466.1
AL391117.1
EFCAB7
AC008574.1
ZNF404
PARD3
TRIQQ
VMP1
MPP6
NMNAT2
NIPAL2
SOGA3
DLG5
MAB21L1
RN7SL67P
POGLUT2
PLS3
F8
AL356968.2
GORC
TMEM14A
SRSF11
TMEM98
GNPDA2
CCDC103
AC002094.1
RASL11A
SSH1
SPTY2D1
RUBCNL
HTRA3
TMEM183B
BCAR1
ZC3H6
CDK2AP1
BX470102.1
AIG1
EPAH2
SLC35F5
LINC01503
TLN1
RIPOR3

Table_S5 genes names positively enriched in MSC

AGKP1
 CHMP1B
 SLC7A8
 ELK3
 SLC4A3
 STAG1
 NAV2
 ZNF334
 ECT2L
 AFAF1
 AFS1
 AC019117.3
 NTF3
 WAC.AS1
 NAV2.IT1
 GCC2
 SCG5
 RAB34
 CNTNAP3P2
 TENM3
 ZMPSTE24
 DMRTA1
 SNORA24
 EPN2
 ANKRD2
 PDK4
 SKAP2
 CSMD2
 BNIP3
 ZSWIM6
 OSTC
 AL109918.1
 CPQ
 ERLEC1
 CDC42BPA
 GPR88
 RBMS3.AS3
 STX7
 MALAT1
 TEK
 GPRASP2
 ANGPTL4
 CPNE7
 WDR35
 CABP1
 BCL10
 WASL
 AC135048.1
 RO60
 SDC4
 PRDM5
 IFT22
 CHIC1
 CHPF
 KIDINS220
 LMCD1
 AGAP1
 FHL1
 ZBTB8A
 AC016747.4
 IGSF10
 GRIA1
 NCS1
 RUFY3
 AC006460.1
 AC064799.2
 ODF3L1
 AC119674.1
 AP005230.1
 BPGM
 MYO1C
 PIP4P2
 RAI2
 CTIF
 TP53BP1
 COL7A1
 AL391422.4
 NOTCH2NLA
 MOCS2
 P2RX6
 PIH1D2
 CCNG1
 LPIN3

Table_S5 genes names positively enriched in MSC

SPATA6
 MFN1
 AC083805.3
 AC090425.1
 SRPK3
 AL157394.1
 AC006213.2
 LRRCS8
 PCDHB10
 AP001528.1
 NPHF3
 HSDL2
 PCDHGA8
 GATA6
 GOLGA4
 AC079601.1
 KIF3A
 KRR1
 HOXA3
 GOLGA8O
 CAMK1G
 PNPLA3
 SINHCAFP3
 TTC26
 PEAR1
 DRAM1
 GADD45A
 MACF1
 IFIT2
 RAP1B
 HOXA6
 SEPT7
 Z99129.4
 ERFE
 UAP1
 ZNF365
 AP002414.2
 TNFAIP8L3
 FNDC4
 TRIM32
 RUSC2
 LAMC2
 MMP3
 PLXNB3
 HSD17B6
 CAMSAP2
 NR3C1
 KCTD9
 C3orf80
 ARHGEF33
 GMNC
 AL590483.1
 CNGA3
 LINC01116
 VSTM5
 CCDC50
 TMEM67
 AC084117.1
 HSPA12A
 FER1L4
 NBR1
 RASSF10
 PSAP
 ZEB1
 IFT88
 CFAP58
 AC006213.5
 HBEGF
 MAOB
 MCHR1
 TMED7
 ARHGAP28
 ZNF300
 AC005736.1
 CMBL
 BCL6
 GAPLINC
 SLC38A5
 SPTBN1
 SCN8A
 SLC22A23
 KDELR2
 MSTN

Table_S5 genes names positively enriched in MSC

APOBEC3C
 AC011558.1
 RHBDP1
 AC135507.1
 KLHL9
 ZNF415
 ARHGAP32
 AL662795.2
 ANKRD12
 ZFYVE9
 SIM2
 CTNNA1
 NUDT6
 LINC00857
 RBPMS
 PAFAH1B2
 SPTBN4
 GOLGA8Q
 GALNT2
 AIDA
 CD276
 NLGN4Y
 RHOF1.AS1
 AC013643.3
 AC036108.2
 TAF13
 FAM92A
 SYNPO
 FBXL2
 CCDC112
 SPATS2L
 ABHD2
 DNAH5
 MAP6
 LINC00842
 RWDD2B
 DUSP1
 CXXC5
 TGM2
 UFSP2
 FAM204A
 KRT222
 TOX2
 GJC1
 RAB7B
 ELN.AS1
 BTBD19
 RASAL2
 STBD1
 MUC1
 GJD3
 CCDC47
 YES1
 MAN1A1
 SNTB1
 PUS7L
 KIF3B
 ZFP36
 AC012370.1
 OMD
 TOR1AIP1
 CACNB3
 TMOD2
 FRMD4A
 UBA5
 MORN4
 BX842568.2
 PCAT6
 SACS
 PCMTD1
 CTDSP2
 AC110792.2
 RBBP4P1
 GLRX
 LINC02605
 GABARAPL1
 NOMO1
 LINC01152
 NFE2L1
 RPL23AP49
 AC119674.2
 ALDH3B1
 ADAMTS15

Table_S5 genes names positively enriched in MSC

EFCAB6
 PURA
 DMBN
 UTP11
 ARFGAP3
 RAPH1
 AC007541.1
 NKIRAS1
 HDGFL3
 ZSWIM8
 TCEAL8
 ZNF423
 BCL2L13
 AC002398.2
 SLC35E4
 MIRLET7BHG
 AADAT
 PALMD
 NR1D1
 CCDC90B
 SH3D19
 LINC00346
 AC000403.1
 SNORA44
 CNTNAP3B
 TSPAN6
 OXTR
 SHOX
 AC102945.2
 LRRCC1
 SNORA40B
 PLXNA1
 ZSCAN26
 NUA1
 RAD50
 JHY
 RPL34P20
 TRPC3
 AL356124.1
 GBAP1
 FTH1P10
 HCFC2
 WBP4
 KAZN
 PAPP2
 MBNL2
 CXCL2
 SLC2A5
 RRAS
 ACTR10
 CADM1
 TRAK2
 CRYZL2P
 SSB
 MEIS3P2
 AP003119.2
 SV2A
 TSPAN1
 ZNF521
 AC090578.1
 AK6
 AL136164.3
 SVIL2P
 STXBP5.AS1
 EXOC1
 FAT2
 CLOCK
 SNORD113.9
 LNP1
 RPS12P16
 IGBP1
 SAV1
 AP000317.1
 MEIS2
 KRT19
 MORF4L1P1
 GNG12.AS1
 APBA1
 LIMS4
 AL117327.1
 TNFRSF10D
 CIZ1
 APOLD1

Table_S5 genes names positively enriched in MSC

COPS2
 RAB3GAP1
 CCDC188
 DUBR
 SH3BP4
 HOXA5
 FARP1
 IL17RD
 MSANTD3.TMEFF1
 NUMB
 MILR1
 STEAP3.AS1
 WFDC1
 SLAIN2
 STAT1
 MDFIC
 FBLN2
 PHTF1
 TANK
 CSRN3P3
 SMPDL3A
 RAB6A
 ATP7A
 TRAF3P2
 FAM160B1
 AL606834.1
 IFI27L2
 KDSR
 RGS20
 LZTFL1
 DDIT4.AS1
 SLC9C1
 RBBP8
 TNFRSF12A
 CLUAP1
 ITGA8
 RNU2.63P
 PREPL
 EXOC5
 NDFIP2
 HBP1
 DSP
 VWA5A
 GCNT1
 AL136116.3
 AC068620.3
 TRIB2
 AC011455.2
 MAFF
 PRKAA2
 DUSP6
 EVI2A
 SEMA6D
 HAPLN1
 BDNF.AS
 AP000941.1
 RNF217
 ZNF703
 TLR4
 JRKL
 LINC01852
 FGF16
 AC026150.1
 DDX43
 TMEM43
 HMGB1P6
 ACTBP7
 AC003101.1
 CCDC91
 ULBP3
 SH3PXD2B
 CDK20
 GJA1P1
 IQCJ.SCHIP1
 AL731568.1
 SAMD9
 KCTD11
 AC008740.1
 C11orf74
 PHF11
 AF165147.1
 C2orf69
 TSPYL2

Table_S5 genes names positively enriched in MSC

VANGL1
AL162231.2
VIM
AC092376.2
PERP
LRRC43
CCDC191
TSC22D2
DLX6.AS1
CDR2L
SNAIL
ZCWPW2
BMP1
HOXA9
ACAD11
TGFB2
RRAGC
AF230666.1
KCNQ5
GRIK1
LAMP2
NFKB1Z
ME1
CNIH1
DEPP1
PDGFD
AC090099.1
EIF4G3
PFN2
L3HYPDH
PPP1R3G
SLFN12
AC022336.3
IL6ST
GIHCG
ZBTB20
ACBD3
HECTD2
KCTD3
CD44
SDC2
LYSMD3
RNLS
AL359091.5
AL035681.1
AHNAK
AL096711.2
NPIP1
DAPK2
FLRT3
EPG5
NUDT12
EFHC2
SAMD9L
TGFB3
DOCK1
STAG2
PNMA8B
CCL2
AC010605.1
AC06249.1
RF00019
EMC4
ANKRD50
PAPSS1
ETFDH
GRB14
PAQR7
GOLGA6L5P
NEK6
NRIP1
PFKM
AC026403.1
IFT140
SMARCE1
NSA2
AL031587.3
SYT7
NOTCH2
PDLIM7
STXBP3
VPS13A
LINC00922

Table_S5 genes names positively enriched in MSC

TBC1D12
LINGO2
STK17B
ADPRH
MAMDC2
AL451064.2
LINC00475
LINC00840
CTBP2P8
AC009404.1
AL391840.2
CABYR
SOX6
NFE2L2
ARHGEF10
CHST1
KANSL1L
FGF18
CSNK1A1
AC005722.2
PKD1P5
SLC12A8
TMEM87B
BMT2
PFDN1
TRIP10
RBP4
SPRY1
R3HDM2
MTUS2
IDI1
AL356515.1
PGF
ACVR1
DDX5
PTP4A2
LIG4
AL355987.2
BMP3
CSTA
MPV17
KCNA4
NOL8
ECHDC1
PCDHB5
LYPD6
PHYHIP
RRAGB
MDGA1
MEG9
LINC00865
SLC9A7P1
ASAP1
LAMB3
FRS2
SDCBP
KYAT3
NMD3
HR
AC012065.2
SLC30A9
AC074143.1
FIG4

Table_S6 summary of qPCR cohort

Type	Number
Adipose MSC	4
BM-MSC	12
Fibroblast of	3
H9 ESC	8
Hepatocytes	3
HUVEC	4
IPS	2
Myoblasts	3
Neural Stem c	2
skin fibroblast	10
Umbilical cor	3

Table_S7 Description of Minc Candidates

target_id	Sleuth qval (adj. Pval)	b (FC estimator)	length (nt)	exons	Chr	Interval	Strand	present in long reads	Genome browser remarks	REFLNC/GTEX	discriminative kmers	expressed in MSCs (kmers)	strong expression in SMCs (kmers)	Notable other expressions	FEELnc_coding_potential
Mlinc.54177.1	5.38147421885392e-10	3.64181797640709	17139	1	18	67485128-67502266	-	FALSE	No annotation (close to predicted lncRNA, overexpression in fibro)	FIBRO	FALSE	Ad/Bm	TRUE	myocytes/osteoblasts	0.336
Mlinc.28428.2	3.91322165116094e-24	4.75090994700694	2816	3	12	88380133-88419354	-	TRUE	TUCP/TCONS_00020862	FIBRO ONLY	TRUE	Ad/Bm	FALSE	dermal papilla/muscle satellites	0.18
Mlinc.109628.1	4.18652552062097e-14	3.5558466039083	3615	2	6	169163268-169168514	+	FALSE	TUCP/TCONS_00012024 (in zone)	FIBRO/ARTERY/BioMarker survie	FALSE	Ad/Bm/UC	TRUE	myocytes/chondrocytes/astrocytes/preadipocytes	0.446
Mlinc.128022.2	2.24440052433889e-15	3.19030456476542	992	5	9	115596550-115692288	+	TRUE	TUCP/TCONS_00016469 (lot of variants in zone) + numerous chromatin marks	FIBRO ONLY	TRUE	Ad/Bm	TRUE	Myocytes/myoblastes	0.268
Mlinc.32708.1	0.00166963056755997	2.11405992530657	4346	1	13	75859960-75864305	+	FALSE	In close 3' region of LOM7	BRAIN (isoform ?)	FALSE	Ad/Bm/UC	TRUE	Chondro (very strong)	0.118
Mlinc.10317.1	8.78516928066226e-11	3.69104237801407	8760	1	1	215239087-215247846	+	FALSE	In close 3' region of KCNK2, very high expression in fibro		TRUE	Ad/Bm	FALSE	osteo/chondro	0.266
Mlinc.45246.1	5.52855779835755e-11	3.12467428011448	3825	1	16	65129303-65133127	+	FALSE	TUCP/TCONS_00024436 (partial overlapping)		FALSE	Ad/Bm/UC	TRUE	SMC (very strong)/osteoblast (very strong)	0.156
Mlinc.61271.1	0.000934727743803686	2.11029387027848	4902	1	2	46810683-46815584	+	FALSE	TUCP/TCONS_00003666, partial overlapping	FIBRO/MUSCLE/ADIPO/Blood	FALSE	Ad/Bm/UC	TRUE	Osteoblast (very strong), chondrocytes (very strong)	0.126
Mlinc.6947.1	0.0087759651506394	1.79363960044882	1126	1	1	148790080-148791205	-	FALSE	On repeated region		FALSE	no expression			0.348
Mlinc.89912.1	2.19410912061805e-13	3.1608578751305	3159	1	4	80263395-80266553	-	TRUE	In close 5' region of FGF5 (very high in fibro)/ Zone of regulation ++		TRUE	Ad/Bm/UC	TRUE	muscle satellite/papilla/myocytes	0.198
Mlinc.98481.1	4.00458831141693e-14	3.43441525297039	1532	1	5	120687803-120689334	+	FALSE	In close 3' region of PRR16 (very high in fibro)		FALSE	Ad/Bm	TRUE	myoblast/chondro	0.084
Mlinc.123973.1	1.61744832508345e-09	2.80103674598808	494	1	9	12823414-12823907	+	FALSE	In close 3' region of LURAP1L (high in fibro)		FALSE	Ad/Bm/UC	TRUE	hepatocytes/osteo (very strong)	0.244
Mlinc.63232.1	1.85061125320116e-11	3.24619016103911	232	1	2	92068044-92068275	-	FALSE	repetition		FALSE	Ad/Bm	TRUE	Myocytes/myoblasts/SMC	0.368
Mlinc.107651.1	1.43285225448328e-06	2.90795380207381	8953	1	6	130445023-130453975	+	FALSE	In close 3' region of TMEM200A (high in fibro)	OVERLAP / PARTIAL	FALSE	Ad/Bm/UC	TRUE	hepatocytes	0.328
Mlinc.117098.1	9.01415915005679e-06	2.68354574301863	1359	1	8	299352-300710	-	FALSE	lincRNA AC136777.1 in zone (10kb)		FALSE	Ad/Bm	FALSE	neuron/muscle/endothelial	0.106
Mlinc.70094.1	5.13172839816861e-10	2.8220665603735	946	1	2	237319841-237320786	-	FALSE	In close 3' region of COL6I3 (very high in fibro)		TRUE	Ad/Bm	FALSE	muscle satellite/papilla/osteo/chondro	0.12
Mlinc.64225.1	0.00206523311314836	2.01363639988553	302	1	21	113609969-113610270	-	TRUE	LTR close to RPL23A? (pseudogene + strong repetition (high in fibro))		FALSE	Bm	TRUE		0.368
Mlinc.54155.1	1.00341251767868e-08	2.89678331657768	5908	1	18	67389037-67394944	-	FALSE	In intron of AC091042.1		FALSE	Ad/Bm	TRUE	osteo/hESC/chondro	0.402
Mlinc.78582.1	1.60651558091348e-09	3.32612797663159	3402	1	3	30418752-30422153	-	FALSE	HMM specific Chr marks (+ FOXA1 site)		NA	NA	NA	NA	0.178
Mlinc.109600.1	1.23389178444261e-08	2.57894271495465	318	1	6	168932197-168932514	+	FALSE	LINE/LTR repetitions? lncRNA in zone		FALSE	Ad/Bm	TRUE	H1-ESC/osteoblast	0.53
Mlinc.9300.1	1.63418031048994e-08	2.66341118521261	1382	1	1	192488616-192489997	-	FALSE	variant of AL390957.1 (Chr marks ++ and FOXP1 site), lot of predicted lncRNA		FALSE	Ad/Bm/UC	TRUE	SMC (very strong)	0.144
Mlinc.54554.1	2.57868369698677e-07	2.45953847799982	2156	1	18	75463916-75466071	+	FALSE	HMM specific Chr marks (+FOXA1 Site)		FALSE	Bm/UC	TRUE	hepatocytes/SMC (very strong)/muscle cell	0.19
Mlinc.83437.1	3.75186606665131e-11	2.92026510167055	2904	1	3	126213050-126215953	+	FALSE	TUCP/TCONS_00006199, lncRNA, Chr Marks		FALSE	Ad/Bm/UC	FALSE	mesenchial/neural prog./myocytes/epithelial	0.292
Mlinc.127918.1	8.59549299747758e-05	2.2148444399395	2870	2	9	113624677-113628441	+	FALSE	LincRNA /TUCP in zone, Chr Mark, TF binding site		TRUE	Ad/Bm	FALSE	Chondro/papilla	0.302
Mlinc.9231.1	2.09587353035184e-06	2.77869731979218	2415	1	1	186681846-186684260	+	FALSE	close to an AS of PTGS2 antisense NFKB1 complex +TF binding sites		FALSE	Ad/Bm/UC	TRUE	Mononuclear cells (very strong)	0.152
Mlinc.78418.1	1.82921400272116e-05	2.52537568052055	2945	1	3	27646119-27649063	+	FALSE	Overlap TUCP/TCONS_i2_00018477 (High in fibro) HMM specific Chr Mark		FALSE	Ad/Bm	FALSE	epithelial/myocytes/chondro	0.232
Mlinc.61267.1	0.00800968729326976	1.85151131179545	8834	1	2	46797931-46906764	+	FALSE	TCONS_00003666, rich zone in lncRNA, Chr mark et TF binding site, close to SOCS5 (high in fibro ++)		TRUE	Ad/Bm	TRUE	osteo/myocytes/papilla	0.162
Mlinc.111761.1	0.000429609804134646	2.16878061892736	1232	1	7	38054085-38055316	+	FALSE	EGRF1 Binding site/ HMM Xhr sites		FALSE	BM	TRUE	Myocytes/myotubes	0.084
Mlinc.54156.1	2.76421814657461e-05	2.55034529045675	6240	1	18	67395327-67401566	-	FALSE	HMM Chr Binding Site + FOXA1 Sites		TRUE	Ad/Bm	TRUE	smc of trachea (very strong)	0.334
Mlinc.47523.1	6.0526557244177e-08	2.53923057783646	1558	1	17	13490975-13492532	-	FALSE	In 3' of gene HS3ST3A1 (High in fibro)	FIBRO/ DIFFUSE	FALSE	Ad	TRUE	epithelial/mesangial	0.342
Mlinc.18093.1	3.3005434599755e-09	2.70413881781352	977	1	11	12085459-12086435	-	FALSE	Close to AC124276.2 variant Chr Mark +++ and TF binding site +++		FALSE	Ad/UC	TRUE	epithelial/keratinocytes	0.236
Mlinc.57835.1	1.15589458808076e-08	2.42525523068851	927	1	19	44737723-44738649	+	FALSE	TUCP/TCONS_00027580		FALSE	FALSE	FALSE	Mononuclear cells (very strong)	0.436
Mlinc.104634.1	7.05793182343196e-12	3.12285461836103	2731	1	6	57260477-57263207	-	FALSE	TUCP/TCONS_00011222 + Chr Mark (HMM) and TF binding site		FALSE	Ad/Bm/UC	TRUE	SMC (strong)	0.314
Mlinc.80687.1	3.47529931298406e-06	2.15786698401882	537	1	3	65018718-65019254	+	FALSE	In Intron of gene ADAM-AS		FALSE	BM	TRUE	epithelial/chondro/neural prog.	0.464
Mlinc.96690.1	2.09413387964091e-06	2.08136277131835	637	2	5	84491916-84493310	+	FALSE			FALSE	BM/UC	FALSE	myocytes/osteo/chondro	0.332

Table_S7 Description of Miinc Candidates

FEELnc label (1-mRNA)	Closest_gene	Distance	TarpmiR max sites by miRNA	Primer1 (F)	Primer2 (R)
0AC110597.1		162		6	
0Y_RNA	-11088		4	GAAGGCCG TGAAGCCAT TGT	CATGATTTT GCCCTCGGC TT
0LINC01615	-344		4		
0U2	42104		4	TGTTTCTGG GCTGCTTGC T	AGGGGACA TTTTCGAC CTCAA
0LMO7	-90		5		
0KCNK2	-1994		5		
0CDH11	-3191		4		
0LINC01118	1084		5		
0AC245389.1	-4592		4		
0FGF5	-46		5	TCGGCGTTG GAAAACCA AAC	TACAGACG GAGAGCTTC CCA
0PRR16	-471		4		
0LURAP1L	-1283		4		
0IGKV1OR2-2	33044		3		
0TMEM200A	-1960		5		
0AC136777.1	-10423		4		
0COL6A3	-3217		4		
0RPL23AP7	-232		4		
0AC091042.1	13732		5		
0LINC01985	-102592		4		
1AL109924.1	30096		4		
0AL136987.1	-27642		4		
0AC116003.2	-26959		5		
0ALDH1L1-AS2	-2881		5		
0AL162727.1	-2043		4		
0PACERR	-400		4		
0AC098614.4	-7343		5		
0LINC01118	9904		5		
0SFRP4	-28390		4		
0AC091042.1	20022		5		
0HS3ST3A1	-3157		4		
0AC124276.2	-456		4		
0BCL3	9056		5		
0RAB23	38163		4		
1ADAMTS9-AS2	-7250		4		
0AC113383.1	-1151		4		

Table_S8 statistically relevant interaction between Mlinc and proteins

Mlinc.28428.2		Mlinc.128022.2		Mlinc.89912.1	
Uniprot ID	gene ID	Uniprot ID	gene ID	Uniprot ID	gene ID
Q496A3	SPATS1	P31994	FCGR2B	P02760	AMBP
Q9UBM1	PEMT	Q9Y2T5	GPR52	Q92830	KAT2A
Q92911	SLC5A5	P14618	PKM	Q6NUI6	CHADL
P09001	MRPL3	Q9P1W8	SIRPG	Q9Y2K9	STXBP5L
Q96G25	MED8	Q9P0J1	PDP1	Q9NV29	TMEM100
P20700	LMNB1	Q8N90	SYT2	Q8N6M3	FITM2
Q5FY80	ARSJ	Q86X59	C17orf82	O43915	VEGFD
O15066	KIF3B	Q8NCL8	TMEM116	Q9NRD1	FBXO6
O77932	DXO	Q9P2T1	GMPR2	Q9UJH8	METRN
Q9P232	CNTN3	O75581	LRP6	Q9BWW8	APOL6
Q9NRY2	INIP	Q7L591	DOK3	Q15746	MYLK
Q12931	TRAP1	Q8IF99	OR10K2	Q9BQE3	TUBA1C
Q0D2K2	KLHL30	Q9UM11	FZR1	Q98HE7	ERO1A
Q16891	IMMT	Q8WUH6	TMEM263	Q8TCG2	PI4K2B
Q9Y6M4	CSNK1G3	Q9UKN8	GTF3C4	Q99969	RARRES2
Q16706	MAN2A1	Q96A47	ISL2	O43148	RNMT
Q99946	PRRT1	Q9ULJ6	ZMIZ1	Q7Z6V5	ADAT2
Q96JJ6	JPH4	Q99829	CPNE1	Q9UIA9	XPO7
Q8IZH2	XRN1	Q29983	MICA	P08684	CYP3A4
Q8NBF1	GLIS1	Q92783	STAM	Q8N135	LG14
Q9BXF6	RAB11FIP5	Q8IYX7	SAXO1	Q96A04	TSACC
Q8TDB4	MGARP	P0DN77	OPN1MW2	Q99217	AMELX
Q9H251	CDH23	P0DN78	OPN1MW3	Q8NAE3	LINC01555
Q96RQ3	MCCC1	Q6ZMH5	SLC39A5	O60403	OR10H2
Q9HBR0	SLC38A10	Q5TH69	ARFGEF3	P49908	SELENOP
Q9BTU6	PI4K2A	P04062	GBA	Q9Y4C5	CHST2
Q9Y2P8	RCL1	Q8IVM8	SLC22A9	Q14142	TRIM14
Q15056	EIF4H	Q96T88	UHRF1	P30085	CMPK1
Q9UMZ2	SYNRG	Q5SWL8	PRAMEF19	G3V0H7	SLCO1B7
Q9GZP1	NRSN2	Q8N427	NME8	P60880	SNAP25
P0C645	OR4E1	Q8NBP7	PCSK9	P01767	IGHV3-53
Q6ZVL6	KIAA1549L	Q8TCX1	DYNC2LI1	Q8WUJ1	CYB5D2
O75909	CCNK	Q13072	BAGE1	Q6ZQN7	SLC04C1
Q99574	SERPIN1	Q86Y27	BAGE5	P83731	RPL24
Q9Y6C7	PCDHGA6	Q96KE9	BTBD6	Q5T1H1	EYS
Q1A5X6	IQCJ	Q6LUX8	CCBE1	P05496	ATP5MC1
Q8IWD5	MFSDBL	Q14409	GK3P	P78312	FAM193A
Q96607	ELF4	P56177	DLX1	P27918	CFP
Q8N6T3	ARFGAP1	Q9H190	SDCBP2	Q15834	CCDC85B
O14950	MYL12B	P23759	PAX7	Q11201	ST3GAL1
Q6ZN28	MACC1	O15409	FOXP2	Q9Y6H1	CHCHD2
O95402	MED26	P60520	GABARAPL2	P68363	TUBA1B
A6NIV6	LRRIQ4	P55809	OXCT1	Q494X3	ZNF404
A8MUM7	LGALS16	Q9NVH2	INTS7	Q96LJ7	DHRS1
P59536	TAS2R41	Q8NOZ8	PUSL1	Q86XS8	RNF130
O95900	TRUB2	Q7L4S7	ARMCX6	P48742	LHX1
Q99593	TBX5	Q9Y5J6	TIMM10B	P15880	RPS2
P21439	ABCB4	Q8WU14	HDAC7	P0DMV0	CT45A7
P20073	ANXA7	P04259	KRT6B	Q8N972	ZNF709
Q8TCC3	MRPL30	Q8IYS8	BOD1L2	Q8IUY3	GRAMD2A
Q6Q788	APOA5	Q95455	TGDS	Q9NR96	TLR9
Q13443	ADAM9	Q969P6	TOP1MT	Q8NBT0	POC1A
Q7L311	ARMCX2	Q3MIP1	ITPR1PL2	O14827	RASGRF2
Q969Q6	PPP2R3C	Q86YV6	MYLK4	Q17R55	FAM187B
Q8N1B4	VPSS2	P04001	OPN1MW	Q9H0M5	ZNF700
P08243	ASNS	Q8NH49	OR4X1	O43715	TRIAP1
Q05193	DNM1	P55771	PAX9	Q9P0V3	SH3BP4
Q13243	SRSF5	Q16656	NRF1	Q04771	ACVR1
Q8NDV7	TNRC6A	Q5JTZ5	C9orf152	Q6ZRS4	ITPRID1
P15941	MUC1	Q95382	MAP3K6	O60337	MARCH6
Q96AN5	TMEM143	Q2QD12	RPEL1	Q13103	SPP2
Q9BX72	CACNG6	Q92796	DLG3	O75663	TIPRL
Q96EY8	MMAB	P05089	ARG1	Q8NBMA	UBAC2
P11498	PC	Q9Y512	SAMM50	Q8NEG0	FAM71C
Q96GC9	VMP1	P20265	POU3F2	O60656	UGT1A9
Q9HC52	CBX8	P04216	THY1	Q7L5L3	GDPD3
Q6UVK1	CSPG4	Q96RY7	IFT140	Q9HV94	RNF122
P08397	HMBS	O43187	IRAK2	Q71U36	TUBA1A
P54922	ADPRH	P21397	MAOA	Q9UJL9	ZFP69B
Q96SA4	SERINC2	Q9H2S9	IKZF4	Q6NUJ1	PSAPL1
Q7L945	ZNF627	Q8N199	ANGPTL6	Q6UXH9	PAMR1
Q8NGB6	OR4M2			P08833	IGFBP1
Q9GZX3	CHST6			O75431	MTX2
P80748	IGLV3-21			Q6UDR6	SPINT4
Q6ZTW0	TPGS1			Q5HYN5	CT45A1
Q8TB05	UBALD1			Q5T0J3	C1orf220
O75947	ATP5PD			Q96T68	SETDB2
O75366	AVIL			Q9UHC3	ASIC3
P32754	HPD			Q96S25	ADO
Q92600	CNOT9			Q8NHU0	CT45A3
Q8TAS1	UHMK1			P29320	EPHA3

Table_S8 statistically relevant interaction between Minc and proteins

P22794	EV12A	Q9H246	C1orf21
A8MTQ0	NOTO	Q9Y5K1	SPO11
Q9Y6U3	SCIN	Q5JVG8	ZNF506
Q96I20	PAWR	Q8NHX4	SPATA3
P06733	ENO1	Q9UM44	HHLA2
P10114	RAP2A	Q9BXM7	PINK1
Q5T4B2	CERCAM	Q99952	PTPN18
Q9HA82	CERS4	A0AVI4	TMEM129
Q8IVT2	MISP	Q9P2R7	SUCLA2
O95427	PIGN	Q9BQB4	SOST
Q9NWB7	IFT57	Q8BX40	LRRRC28
P0CC24	NJDT7	Q8N7S2	DNAJC5G
Q9Y241	HIGD1A	Q9BPO6	ZNF317
P55291	CDH15	Q9P2U8	SLC17A6
Q3ISF7	ACOT6	Q4G0U5	CFAP221
Q9H467	CUEDC2	Q9H9B1	EHMT1
Q86XN7	PROSER1	Q0P670	SPEM2
Q07075	ENPEP	O94806	PRKD3
Q5JPH6	EARS2	O75695	RP2
Q6ZU45	CLEC20A	Q3MJ13	WDR72
Q9Y263	PLAA	A8MXK1	VSTM5
O75309	CDH16	Q96NJ6	ZFP3
Q8N608	DPP10	Q6J9G0	STYK1
P36871	PGM1	Q9Y5L0	TNPO3
Q8WXF5	CRYGN	Q5TAT6	COL13A1
Q8IYX0	ZNF679	P59646	FXYD4
Q8NDC0	MAPK1IP1L	Q6UXZ4	UNC5D
Q68DC2	ANKS6	P09565	GIG44
Q9NRA2	SLC17A5	P0CG47	UBB
Q5VZE5	NAA35	Q14CZ8	HEPACAM
P13521	SCG2	Q86US8	SMG6
Q9BU23	LMF2	Q9Y6I4	USP3
P48382	RFX5	Q9BFO7	LINC00526
P35222	CTNBNB1	Q07654	TFF3
Q7Z7M9	GALNT5	Q9BRD2	OR52B2
Q8WXD9	CASKIN1	Q8WWQ8	STAB2
P04843	RPN1	P78545	ELF3
Q9BYD1	MRPL13	Q96M02	C10orf90
Q9Y573	IPP	Q9NRY7	PLSCR2
O00748	CES2	Q12767	TMEM94
Q9UK45	LSM7	Q96920	TBRG4
Q8NFJ9	BBS1	P35410	MAS1L
O43525	KCNQ3	P50052	AGTR2
Q96T83	SLC9A7	Q6VY07	PACS1
O95104	SCAF4	Q8NH04	OR2T27
Q9HBJ1	PARVB	Q70EL2	USP45
Q9P2V4	LRIT1	P62244	RPS15A
Q9HCN3	TMEM8A	Q9H2A7	CXCL16
P11464	PSG1	Q9BT92	TCHP
Q9UNN4	GTF2A1L	P16455	MGMT
O75935	DCTN3	P0DMU7	CT45A6
Q6V0L0	CYP26C1	P0DMU8	CT45A5
Q9H488	POFUT1	P81277	PRLH
Q9NRM1	ENAM	Q14192	FHL2
Q9NYP3	DONSON	Q15486	GUSBP1
O60636	TSPAN2	Q8IYB5	SMAP1
Q86Y28	BAGE4	Q12791	KCNMA1
Q6P4R8	NFRKB	Q8NGT1	OR2K2
O95319	CELF2	O00755	WN17A
Q9Y585	OR1A2	P23290	CA6
P20339	RAB5A	Q7RTS6	OTOP2
Q9UBI4	STOML1	P10720	PF4V1
A4D1S5	RAB19	Q9H553	ALG2
P51523	ZNF84	Q86YD1	PTOV1
Q96I99	SUCLG2	Q27J81	INF2
Q9BR39	JPH2	Q86Y38	XYLT1
Q14032	BAAT	O95363	FARS2
Q9COK1	SLC39A8	A8MV57	MPTX1
Q9UQ16	DNM3	Q9H6P5	TASP1
P78344	EIF4G2	Q15084	PDIA6
Q1ZZU3	SWI5	Q8N690	DEFB119
P55157	MTTP	Q96JN8	NEURL4
Q5TBB1	RNASEH2B	Q9H3R0	KDM4C
O75175	CNOT3	Q96E17	RAB3C
Q14558	PRPSAP1		
P08183	ABCB1		
P13498	CYBA		
Q9BRB3	PIGQ		
Q96C34	RUNDC1		
Q9P0K8	FOXJ2		
Q15034	HERC3		
Q9UI30	TRMT112		
Q9NUQ7	UFSP2		

Table_S8 statistically relevant interaction between Minc and proteins

Q9HAH1	ZNF556
Q8NHS0	DNAJB8
Q9BWM5	ZNF416
Q32P51	HNRNPA1L2
Q8NFT6	DBF4B
Q8N1N5	CRIPAK
Q9Y5F6	PCDHGC5
P50454	SERPINH1
P08247	SYP
O14979	HNRNPD1
Q8NB16	XXYL1
Q13885	TUBB2A
Q9NQS3	NECTIN3
Q8NGU4	OR21P
Q9NX63	CHCHD3
O60907	TBL1X
P08697	SERPINF2
P52738	ZNF140
Q14406	CSHL1
P62333	PSMC6
Q9BTV5	FSD1
Q6U7Q0	ZNF322
Q05D32	CTDSPL2
Q9H857	NT5DC2
P19105	MYL12A
A6NML5	TMEM212
Q9NSI2	FAM207A
Q6IC98	GRAMD4
Q96IW7	SEC22A
Q96HL8	SH3YL1
Q8IUR5	TMTC1
Q8N5R6	CCDC33
P06732	CKM
P55089	UCN
Q8N2G8	GHDG
Q96H40	ZNF486
Q6P3S6	FBXO42
P56282	POLE2
Q8TE12	LMX1A
Q9NPR9	GPR108
P54709	ATP1B3
O95498	VNN2
Q969V3	NCLN
Q15398	DLGAP5
P43686	PSMC4
Q14118	DAG1
O43414	ERI3
P40225	THPO
A0AV96	RBM47
Q9HCM3	KIAA1549
Q5SR56	MFS14B
Q96MG7	NSMCE3
Q96M29	TEKT5
C9JDP6	CLDN25
Q99674	CGREF1
Q5VZ18	SHE
Q8IYU8	MICU2
Q9UNZ2	NSFL1C
Q9BX68	HINT2
P63124	HERV-K104
Q8NGH7	OR52L1
Q9UMD9	COL17A1
Q02548	PAX5
A2RU54	HMX2
O43570	CA12
O60928	KCNJ13
Q9BRQ8	AIFM2
Q96M93	ADAD1
Q6PI98	INO80C
Q9Y3B7	MRPL11
Q13310	PABPC4
A6NKC9	SH2D7
P20248	CCNA2
O95817	BAG3
O75787	ATP6AP2
A6N115	MSGN1
Q15653	NFKB1B
Q6QNY0	BLOC1S3
Q6ZMV8	ZNF730
Q9NP70	AMBN
Q86UY5	FAM83A
Q5DX21	IGSF11
Q9NYS0	NKIRAS1

Table_S8 statistically relevant interaction between Mlinc and proteins

Q8NG35	DEFB105A
Q8NG35	DEFB105B
Q9Y3C7	MED31
Q14126	DSG2
P10745	RBP3
Q15125	EBP
Q969G6	RFK
Q96115	SCLY
Q2VPB7	AP5B1
Q8IXB3	TRARG1
Q6VX0	CYP2R1
Q96M83	CCDC7
Q96QD9	FYTTD1
Q9NRK6	ABCB10
Q7L112	SV2B
Q86YD3	TMEM25
Q16557	PSG3
P27105	STOM
Q81ZD4	DCP1B
Q8NE01	CNNM3
Q30KQ4	DEFB116
O95873	C6orf47
Q8NAN2	MIGA1
Q9H1Q7	PCED1A
A6NHNO	OTOL1
P59091	LINC00315
Q9H9V9	JMJD4
Q9HBL0	TNS1
P51397	DAP
Q07157	TJP1
Q8IY34	SLC15A3
P30989	NTSR1
Q8NGP4	OR5M3
Q96R84	OR1F2P
Q572E6	ARMH3
Q86TN4	TRPT1
Mlinc.128022.2	
Uniprot ID	gene ID
P31994	FCGR2B
Q9Y2T5	GPR52
P14618	PKM
Q9P1W8	SIRPG
Q9P0J1	PDP1
Q8N9I0	SYT2
Q86X59	C17orf82
Q8NCL8	TMEM116
Q9P2T1	GMPR2
O75581	LRP6
Q7L591	DOK3
Q6IF99	OR10K2
Q9UM11	FZR1
Q8WUH6	TMEM263
Q9UKN8	GTF3C4
Q96A47	ISL2
Q9ULJ6	ZMIZ1
Q99829	CPNE1
Q29983	MICA
Q92783	STAM
Q8IXX7	SAXO1
P0DN77	OPN1MW2
P0DN78	OPN1MW3
Q6ZMH5	SLC39A5
Q5TH69	ARFGEF3
P04062	GBA
Q8IVM8	SLC22A9
Q96T88	UHRF1
Q5SWL8	PRAMEF19
Q8N427	NME8
Q8NBP7	PCSK9
Q8TCX1	DYNC2L1
Q13072	BAGE
Q86Y27	BAGE5
Q96KE9	BTBD6
Q6UXH8	CCBE1
Q14409	GK3P
P56177	DLX1
Q9H190	SDCBP2
P23759	PAX7
O15409	FOXP2
P60520	GABARAPL2
P55809	OXCT1
Q9NVH2	INTS7
Q8N0Z8	PUSL1

Table_S8 statistically relevant interaction between Mlinc and proteins

Q7L4S7	ARMCX6
Q9Y5J6	TIMM10B
Q8WUI4	HDAC7
P04259	KRT6B
Q8IYS8	BOD1L2
O95455	TGDS
Q969P6	TOP1MT
Q3MIP1	ITPRIPL2
Q86YV6	MYLK4
P04001	OPN1MW
Q8NH49	OR4X1
P55771	PAX9
Q16656	NRF1
Q5J7Z5	C9orf152
O95382	MAP3K6
Q2QD12	RPEL1
Q92796	DLG3
P05089	ARG1
Q9Y512	SAMM50
P20265	POU3F2
P04216	THY1
Q96RY7	IFT140
O43187	IRAK2
P21397	MAOA
Q9H2S9	IKZF4
Q8NI99	ANGPTL6