Human Reproduction, 2016
doi:10.1093/humrep/dew008

human reproduction SUPPLEMENTARY DATA

Supplementary methods

Full description of laboratory and bioinformatics pipeline.

Human endometrial cell culture

The single-cell suspension (100 μl) from the previous step was placed in cell culture medium consisting of a 1:1 mixture of DMEM (Cat No E15-892, GE Healthcare, US)/Ham's F12 (Cat No E15-890, GE Healthcare), supplemented with 10% FBS and Antibiotic-Antimycotic solution (GE Healthcare). After 8 days the cultured cells were rinsed twice with PBS and dissociated with 0.25% Trypsin-EDTA (Life Technologies, USA). After trypsin treatment, the cell culture medium was added, cells were collected and centrifuged at 205 × *g* 4°C for 6 min. The cell pellet was suspended in 4 ml ice-cold PBS containing 5% FBS solution and the suspension was filtered twice through 50 and 35 μm Falcon Tubes with Cell Strainer Cap (Fisher Scientific, US). The filtrate was centrifuged at 210 × *g* 4°C for 6 min to collect cells and re-suspend them in 100 μl of PBS/FBS solution. Endometrial stromal single cells were stained with fluorescence-conjugated antibodies and handled as described in the previous section.

Single-cell flow cytometry

In order to maintain a high viability ratio, all handling procedures were performed on ice and the cytometer sample chamber was pre-cooled to 4°C. BD FACSAria I cell sorter (BD Biosciences, USA) under a pressure of 20 PSI, equipped with a 100 μm nozzle, was used. Prior to sorting, possible cell aggregates were removed by filtering the cell suspension through a 35 μm nylon mesh. Filtered and autoclaved PBS solution was used as sheath fluid for all flow cytometry experiments. FITC and R-Phycoerythrin fluorescent dyes were excited by a blue 488 nm laser and detected using 530/30 and 585/42 optical filters, respectively. DAPI was excited using a UV 355 nm laser and detected using a 450/40 optical filter. During sorting, the 'sweet spot' button was enabled and 'single-cell' was selected as the sorting precision mode. Minimum 10 000 events were acquired for each sample. After compensation, the matrices were updated and sorting gates adjusted according to the obtained data. The cellular populations representing stromal and epithelial cells generally appeared on the same place on the plot, however minor differences between samples were still present. Cells were sorted into a half 96-well PCR plate. A custom adapter, ~50% higher than the standard one, was used to fix the 96-well PCR plate within the FACS ARIA automated cell deposition unit. By introducing this adapter the flight length of drops/cells was shortened by about 15 mm, resulting in higher single-cell sorting accuracy. Data analysis was performed using the FACS DiVa software (BD Biosciences).

Cells were sorted directly into 5 μl of mild hypotonic lysis buffer containing 25 mM Tris–HCl (pH 8.0) (Sigma-Aldrich (Sigma), Germany), 2 mM dNTP mixture (Thermo Fisher Scientific (Thermo), USA), 0.7 U/μl RiboLock RNase Inhibitor (Thermo), 0.2% Triton X-100 (Sigma), 400 nM anchored oligo-T 5′-Biotin-TTAAGCAGTGGTAT CAACGCAGAGTCGACT$_{30}$V-3′, where 'V' is either 'A', 'C' or 'G' and locked nucleic acid nucleotide (Exiqon, Denmark), and 2 μM barcoded template switching oligonucleotide (Sigma, Supplementary Table SI). Non-skirted low profile 96-well plates (Thermo) were used with proper adhesive PCR film (Cat No AB-0558, Thermo). Already sorted plates were sealed immediately, snap-frozen on dry ice and stored at −80°C until further use. In general, sorted cells were exposed to room temperature (RT) for less than 2 min, minimising possible RNA degradation due to elevated temperatures.

Reverse transcription

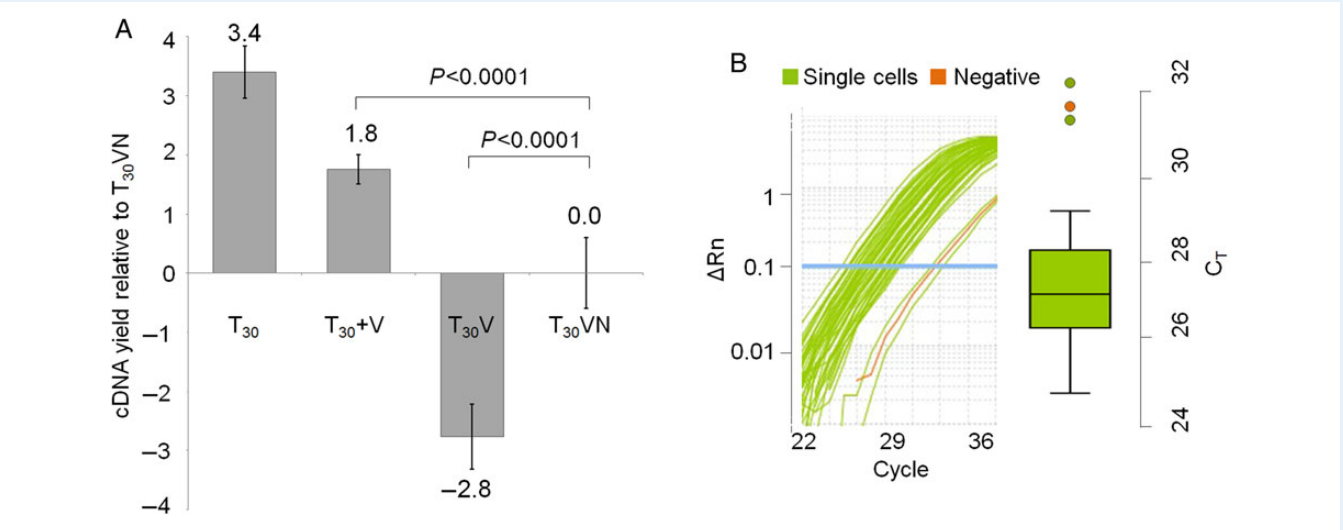
The FACS sorted plate was placed directly from −80°C onto a pre-heated thermoblock (T100, Bio-Rad Laboratories, USA) at 75°C for 3 min and immediately placed on ice afterwards. Five microlitres of first-strand mastermix, containing 100 mM Tris–HCl (pH 8.0) (Sigma), 20 mM $MgCl_2$ (Sigma), 150 mM KCl (Sigma), 2 M Betaine (Sigma), 10 mM DTT (Invitrogen, USA), 1.5 U/μl RiboLock RNase Inhibitor (Thermo), 14 U/μl RevertAid Premium reverse transcriptase (Thermo) and 1:10 000 000 of ERCC RNA Spike-In Mix I (Life Technologies) was added to each lysed cell at 4°C thermoblock and mixed by pipetting 10 times. One microlitre of diluted (1:10 000 000) Spike-In Mix I was added to the whole 48-cell library mastermix in a 280 μl volume. Finally, the reverse transcription reaction was carried out by incubating at 42°C for 90 min, followed at 85°C for 5 min for enzyme inactivation.

Single-cell sorting quality control

From the 10 μl of cDNA, 1 μl was used for quality control (QC) of the single-cell FACS accuracy and cDNA synthesis. After cDNA synthesis, 1 μl of the product was added to 19 μl of qPCR reaction mixture containing 4 μl 5× Hot FirePol EvaGreen qPCR Mix Plus with ROX (Solis BioDyne, Estonia) and 500 nM universal primer 5′-AAGCAGTGGT ATCAACGCAGAGT-3′ (Sigma) (Supplementary Fig. S1). The reaction was incubated at 95°C for 15 min and then cycled 40 times between 95°C 15 s, 63°C 15 s and 72°C 1 min using the standard ramp speed of the 7500 Fast Real-Time PCR System (Life Technologies).

Oligo-T comparison

As reverse transcription is a key step for single-cell gene expression studies, we performed an oligo-T comparison to evaluate the most efficient oligo-T anchored primer. The previously here described reverse transcription and qPCR-based quantification were followed with minor modifications. Ten picograms of human endometrial RNA in one microlitre were used instead of a single-cell, and different oligo-T primers



Supplementary Figure S1 Oligo-T primer comparison and single-cell FACS quality control. (**A**) Average yield of synthesized cDNA using different oligo-T primers, relative to those obtained using the oligo-T_{30}VN oligo. Full sequences are provided in Supplementary Table SI. (**B**) Single secretory biopsy stromal cells were lysed and template-switching mediated reverse transcription was used to generate cDNA. Of the total volume of single-cell cDNA, 10% was used for FACS accuracy and cDNA synthesis QC. A typical QC layout showing that the RNA from the majority of sorted cells is converted successfully to cDNA. The negative control had a significantly higher C_T value and clustered with two samples, suggesting that the cells in those wells most likely did not reach to the lysis buffer during the FACS procedure. The negative C_T value would have been even higher if artificial spike-in molecules had not been used in the mastermix.

(Supplementary Table SI and Supplementary Fig. S1) were analysed simultaneously. For qPCR, 10-fold dilutions were created to convert relative C_T value to absolute fold changes.

cDNA capture and amplification

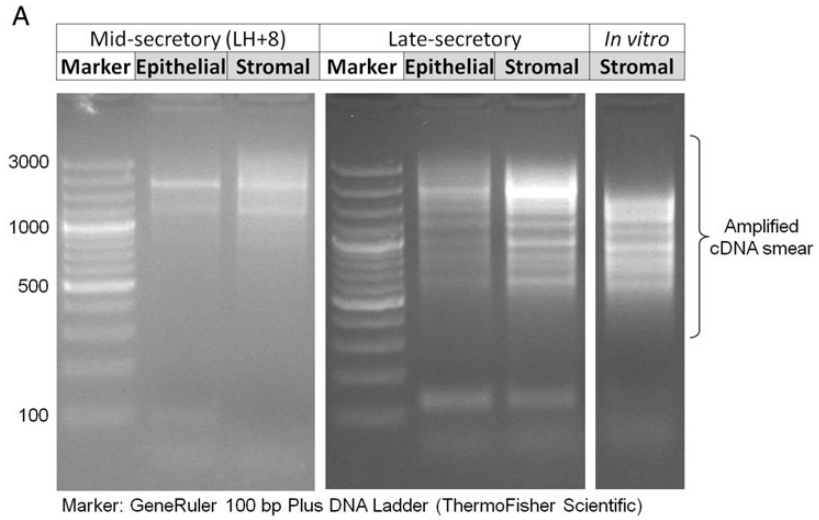
The first-strand product from individual 48-wells was captured, size-selected and concentrated by carboxylated magnetic beads. For one 48-cell library, 50 µl Dynabeads MyOne Carboxylic Acid (Life Technologies) was mixed with 400 µl capture buffer, containing 25% PEG-6000 (Sigma), 2.25 M KCl, and 10 mM Tris–HCl (pH 7.5) (Sigma). A multichannel pipette was used to collect 48-well reactions into a 2 ml DNA low-bind tube (Eppendorf, Germany) together with 450 µl of capture buffer. The volume of the mixture was measured and nuclease-free water was added up to 1 ml to ensure a final 0.9 M KCl and 10% PEG-6000 for cDNA size-selection and concentration. The standard 0.9 M NaCl in binding buffer was replaced by 0.9 M KCl to skip further beads ethanol washing step and avoid Na^+ negative influence to following Phusion (Thermo) amplification. By contrast, additional K^+ ions increase the Phusion's PCR yield, eliminating the need of ethanol washing.

After 20 min of incubation at RT, beads were captured for 3 min by DynaMag-2 (Life Technologies). The supernatant was removed carefully, after this the beads were spun for 10 s and placed back on the magnet. The remaining capture buffer was removed with the pipette and the beads were directly (without ethanol washing) resuspended in 75 µl elution buffer (EB, 5 mM Tris–HCl (pH 8.5)) (Sigma) and 0.02% Tween-20 (Sigma). After 3 min at RT, beads were captured by the magnet and 75 µl of supernatant was used as a template for first-strand amplification. A PCR cycler (Bio-Rad) was preheated to 98°C and 25 µl of first-strand PCR mastermix containing 1× Phusion GC buffer (Thermo), 250 µM dNTP mixture (Thermo), 2 µM universal primer 5′-Biotin-TEG-AAGCAGTGGTATCAACGCAGAGT-3′ (Sigma) and 4 U Phusion High-Fidelity polymerase (Thermo) were added. The reaction was incubated at 98°C for 2 min followed by 25 cycles of 98°C for 15 s, 65°C for 1 min and 72°C for 5 min with a final extension at 72°C for 5 min using the default ramp speed of the T100 cycler (Bio-Rad).

Library preparation

Thirty microlitres of Dynabeads MyOne Streptavidin C1 (Life Technologies) were washed twice with 2× biotin-streptavidin binding buffer (10 mM Tris–HCl (pH 7.5) (Sigma)), 1 mM EDTA (Sigma), 2 M NaCl (Sigma), and 0.01% Tween-20 (Sigma) on the magnet stand (DynaMag-2, Life Technologies) and eluted finally in 100 µl 2× buffer. The first-strand PCR product (100 µl) was mixed with washed beads and incubated for 20 min on a 400 rpm shaker at RT. The beads were washed twice with EB and finally resuspended in 40 µl EB. Enzymatic fragmentation was performed in a 50 µl volume containing 1× Fragmentase buffer V2 and 5 µl Fragmentase enzyme (both from New England Biolabs, US (NEB)). The reaction was incubated for 30 min on a 400 rpm shaker at 37°C, terminated by double EB wash and subsequently eluted in 40 µl EB. Fragment end-repair was carried out in a 50 µl volume containing 1× NEBNext End Repair Buffer, 5 µl NEBNext End Repair Enzyme Mix, and 1 µl *E. coli* DNA ligase for Fragmentase (all from NEB). The mixture was incubated for 30 min on a 400 rpm shaker at 20°C, terminated by a double EB wash and lastly eluted in 42 µl of EB. Five microlitres of 10× NEBNext dA tailing buffer and 3 µl of Klenow exo- (5 U/µl) (both from NEB) were added and incubated for 30 min on a 400 rpm shaker at 37°C. The reaction was stopped by two consecutive EB washes and the beads were then

A



Marker: GeneRuler 100 bp Plus DNA Ladder (ThermoFisher Scientific)

B

	Mid-secretory (LH+8)		Late-secretory		*In vitro*
	Epithelial	**Stromal**	**Epithelial**	**Stromal**	**Stroma**
Raw reads per cell	3 804 806	3 996 306	3 046 911	2 868 949	3 224 877
Low quality read rate (%)	10.14	13.60	11.37	12.69	13.01
Improper STRT read rate (%)	10.86	9.43	9.70	9.69	8.80
PCR redundancy	**144.48**	**81.43**	**57.27**	**41.16**	**38.82**
Qualified reads per cell	20 854	36 868	38 840	53 326	63 590
Mapped rate (%)	56.85	73.68	60.93	77.14	82.52
Mapped reads per cell (median)	**6 330**	**21 486**	**7 223**	**37 298**	**55 974**
Spike-in reads per cell	1 897	1 268	540	305	151
5'-end capture rate (spike-in, %)	97.69	99.18	97.23	96.36	96.65
5'-end capture rate (mRNA, %)	57.02	40.53	72.84	70.69	86.36

Supplementary Figure S2 Single-cell quality control on agarose gel and by RNA-seq. (**A**) Pooled single-cell cDNAs are amplified by PCR and loaded on 2% agarose TAE gel for QC. Equal cDNA broad smear is expected from 0.2–5 Kb range. (**B**) Epithelial and stromal cells from two biopsies were sequenced on one Illumina HiSeq2000 lane. Stromal cells from late secretory biopsy were cultivated and analysed in parallel (pointed by arrow). The description of all measured RNA-seq QC parameters are described in https://github.com/shka/STRTprep.

eluted into 40 μl of EB. Twenty units of High Fidelity SalI restrictase were used in a 50 μl reaction volume together with 1 × CutSmart Buffer and 1 mM ATP (all from NEB) to eliminate a transcript's 3′ end universal sequence and to reduce reads from that end. The reaction was next incubated for 30 min on a 400 rpm shaker at 37°C, followed by two EB washes. Finally, beads were eluted in 18 μl of EB.

A double-stranded adapter was formed after heating both oligonucleotides (5′-CAAGCAGAAGACGGCATACGAGCTCTTCCGATCT-3′ and 5′-Pho- GTTCGTCTTCTGCCGTATGCTCGAGAAGGCT AG-3′-Pho) at 95°C and cooled slowly to RT in a solution with 10 mM Tris–HCl (pH 7.6) and 100 mM NaCl. Later, 3 μl of T4 DNA ligase 10× buffer (Thermo) was added to the previously eluted 18 μl SalI-restricted product. In addition, 3 μl of 50% PEG-6000, 3 μl of the double-stranded adapter, and 3 μl of T4 DNA Ligase (5 U/μl) (Thermo) were added. The mixture was incubated for 30 min at 21°C without shaking and was pipetted during this time. The ligation was terminated by two consecutive EB washes and two water washes. Beads were re-suspended in 75 μl of water and incubated at 75°C for 3 min to break the biotin-streptavidin bonds and allow the release of DNA molecules for further amplification. Beads were captured by a magnet and the supernatant was transferred directly to a new PCR tube.

For the second round of PCR in which Illumina cluster generation sequences are introduced, 20 μl of Phusion HF 5× buffer (Thermo) was added to the previous 75 μl product. In addition, primers (5′-AA TGATACGGCGACCACCGAGATCTAAGCAGTGGTATCAACGCA GAGT-3′ and 5′-CAAGCAGAAGACGGCATACGAG-3′) at final 500 nM concentration, dNTPs at 200 μM, and 2 U of Phusion polymerase (Thermo) were added to a final PCR volume of 100 μl. The reaction was incubated at 98°C for 1 min, followed by 18 cycles of 98°C for 15 s, 65°C for 30 s and 72°C for 30 s with a final extension at 72°C for 1 min using the default ramp speed of the T100 cycler (Bio-Rad).

Library size-selection

The library was size-selected using AMPure XP beads (Beckman Coulter, US) in a two-step selection. First, 70 μl of beads were added to 100 μl of the PCR product (ratio 0.7 ×), incubated for 5 min at RT and captured by a magnet for 3 min. One hundred sixty microlitres of supernatant were transferred to a clean tube and 35 μl of beads (0.22 ×) were added. The mixture was incubated for 5 min at RT and beads were captured in the magnet stand for 3 min. After incubation, the supernatant was discarded and the remaining beads were centrifuged at 500 × *g* for 10 s. After centrifugation, the beads were placed again on the magnet and all remaining supernatant was removed. The beads were eluted directly without ethanol washing in 25 μl of EB and incubated for 1 min at RT. Finally, the eluted library was transferred to a clean tube after 3 min incubation on the magnet. The library was visualized on a TapeStation High Sensitivity D1000 ScreenTape (Agilent Technologies, USA) and quantified using the KAPA Library Quantification Kit (KK4835, KAPA Biosystems, South Africa).

Cluster generation and sequencing

Cluster formation and sequencing-by-synthesis was performed on an Illumina cBot and HiSeq2000 instrument (Illumina, USA) following an in-house low-concentration protocol from 0.15 nM. Up to 8 μl of sample was used to achieve a final concentration of 10 pM in 115 μl hybridization buffer. One microlitre of 2 M NaOH was added and incubated for 5 min at RT to denature the DNA. The reaction was neutralized by adding 1 μl of 2 M HCl and 105 μl of HT1 buffer (Illumina). A single-read 59 bp run with a custom primer (5′-GATCTAAGCA GTGGTATCAACGCAGAGT-3′) at 600 nM concentration was used to get UMI information, the sample barcode, and expression data as described previously (Islam *et al.*, 2012, 2014).

Bioinformatics

Preprocessing

The raw sequenced reads were trimmed to a length of 50 bp. Low-quality reads, with at least one low-quality base within the trimmed region, were excluded. Redundant reads in the qualified ones were collapsed into the best-quality read. The collapsed reads were demultiplexed by comparison with the 48 barcodes (Supplementary Table SI) plus the template switching adaptor sequences; non-canonical collapsed reads which did not match any sequences within a less than two Levenshtein distance, were excluded. Finally, the prefixes for barcoding (6 bp) and template switching guanines (3 bp) in the canonical reads were trimmed.

Alignment

The trimmed canonical reads were aligned by TopHat (Kim *et al.*, 2013) with Bowtie (Langmead *et al.*, 2009), a genome index, and a transcriptome index. The genome index consists of the UCSC hg19 (Lander *et al.*, 2001), ERCC synthetic spike-in RNAs (Life Technologies), the synthetic construct spike-in microarray control hypothetical protein (ynbA) gene, partial CDS (GenBank: EF011072) and human ribosomal DNA complete repeating unit (GenBank: U13369). The transcriptome index consists of RefSeq (Pruitt *et al.*, 2005) alignments in UCSC Genome Browser annotations (Karolchik *et al.*, 2014). Both of the indexes were built by scripts in STRTprep. Alignments with the same reads mapping to multiple loci were excluded.

Quality checking

Values of (i) $\log_{10}$ of spike-in read counts, (ii) $\log_{10}$ ratios of total mapped read counts versus spike-in read counts, (iii) 5′-end capture rates of spike-in RNAs, and (iv) 5′-end capture rates of protein-coding genes were calculated for each sample, and the distributions were summarized per library. Outlier samples in the distributions, defined as extreme values outside the Q1-1.5·interquartile range (IQR) or Q3+1.5·IQR, were excluded from further analysis.

Quantification and normalization

The uniquely aligned reads within the 5′-UTR of protein-coding genes or 500 bp upstream were counted by genes. In addition, reads within 50 bp at the most 5′-end were counted by spike-in RNAs. The read counts were summarized for each qualifying sample. Normalized expression levels were calculated as relative levels versus the sum of spike-in reads in each sample.

Estimation of fluctuation

Fluctuation is the degree of variation of the normalized expression levels versus an expected level of technical noise. Significant fluctuation means that the degree of variation is significantly greater than the noise level. The expectation for target samples in a target library was estimated by fitting a generalized linear model with a Gamma distribution and identity link to $y = a_1/x + a_0$, where x is the average, and y is the squared coefficient of variation of the normalized spike-in levels. Integrated expectation over the target libraries was estimated by summation and scaling properties on a Gamma distribution. Fluctuation and significance were evaluated by using an integrated noise model. A detailed explanation of fluctuation is provided in Supplementary Text SI.

Estimation of differential expression

Differential expression is the degree of difference of normalized expression levels between predefined classes of the target samples, assessed by SAMstrt (Katayama *et al.*, 2013), while fluctuation is independent of classification. To use block permutation in a multi-class comparison, we used SAMseq (Li and Tibshirani, 2013) with a modification as the back-end (https://github.com/shka/samr/tree/test_multblock).

Heatmap

The heatmap is produced by Heatplus (Ploner, 2014), using the Spearman rank correlation distance and the Ward clustering method versus the $\log_{10}$-normalized expression levels of significantly fluctuating (adjusted P-value < 0.05) and differentially expressed (q-value < 0.05) genes in the qualifying target samples, with the addition of the minimum but non-zero normalized value in the expression profile to avoid $\log_{10}(0)$.

Gene ontology analysis

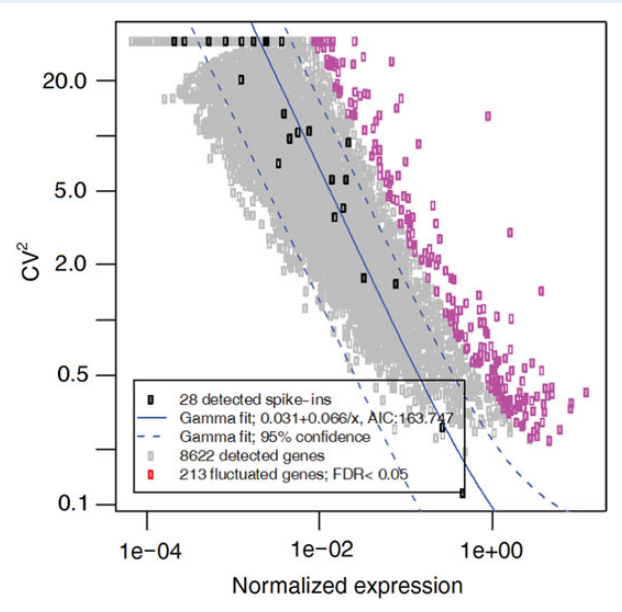
Over-represented functional annotations in the differentially expressed genes were detected by using the functional tool FatiGO, which is available in the *Babelomics* online platform (Medina *et al.*, 2010).

Supplementary Text SI The concept of 'fluctuation'

SAMseq (Li and Tibshirani, 2013) and SAMstrt (Katayama *et al.*, 2013) estimate the significance of differential expression with respect for technical noise, which is the variation in sequencing depths and the sampling noise. SAMstrt uses only the total counts of the spike-in RNAs as the depths (d_i in Li *et al.* (Li and Tibshirani, 2013)) for the spike-in based normalization purpose, while SAMseq uses the total counts of all genes. This means that the noise component in the observed spike-in read counts affects the level of the normalized values. Therefore, to be of interest, the fold difference or the degree of differentially expressed genes, should be larger than the technical noise level.

'Fluctuation' is one solution to estimating the degree of variation. We used several spike-in RNAs, added to all samples at the same concentration, for estimation of the noise level. In detail, expectation of the noise is modelled by fitting against a Gamma distribution and identity link to $y = a_1/x + a_0$, where x is the average, and y is the squared coefficient of variation (CV^2), of the normalized spike-in levels (black points and blue lines in Supplementary Fig. S3). We can then find the significance of the degree of variation in each gene by comparison to the expected noise level (magenta points in Supplementary Fig. S3). For robustness against library biases, the noise modelling is performed by library, and the noise models and the observed CV^2 over all target libraries are integrated by summation and scaling properties on the Gamma distribution. Finally, the significance of the integrated variation in each gene is evaluated by comparison to the integrated noise model; the fluctuationScore in out/byGene/diffexp.xls is $CV^2_{observed}/CV^2_{expected}$, and fluctuation is the adjusted *P*-value.

Further thresholding of differentially expressed genes by the significant fluctuation reduces false positives from the technical noise in the spike-in read counts. The fluctuation estimation does not need a classification of the samples, while the differential expression test requires predefined classifications. You can therefore use this measure to select highly 'variable' genes within a cell population as (Brennecke *et al.*, 2013), or between the patients, and so on.



Supplementary Figure S3 Estimation of fluctuation among cultured cells.

Supplementary Text SII

DEFAULTS: &defaults

UMI: 4
BARCODE: 6
GAP: 3
CDNA: 37
LAYOUT: src/barcodes.old.txt
PHYX:
GENOMESPIKERIBO: src/ebwt/hg19_ercc92_ynbA_u13369/ref
TRANSCRIPT: src/ebwt/hg19_refGene/ref
FLUCTUATION: 0.05
DIFFEXP: 0.05

LIBS:

SARMB:

<<: *defaults
FASTQS:

- src/150221_SN670_0304_AC655YACXX-lane7_NoIndex_L007_R1_001.fastq.gz
- src/140528_SN653_0292_AC4WG4ACXX-lane6_NoIndex_L006_R1_001.fastq.gz

SARMC:

<<: *defaults
FASTQS:

- src/150221_SN670_0304_AC655YACXX-lane8_NoIndex_L008_R1_001.fastq.gz
- src/140528_SN653_0292_AC4WG4ACXX-lane8_NoIndex_L008_R1_001.fastq.gz

References

Brennecke P, Anders S, Kim JK, Kolodziejczyk AA, Zhang X, Proserpio V, Baying B, Benes V, Teichmann SA, Marioni JC *et al.* Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods* 2013; **10**:1093–1095.

Islam S, Kjallquist U, Moliner A, Zajac P, Fan JB, Lonnerberg P, Linnarsson S. Highly multiplexed and strand-specific single-cell RNA 5′ end sequencing. *Nat Protoc* 2012;**7**:813–828.

Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, Lonnerberg P, Linnarsson S. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods* 2014;**11**:163–166.

Karolchik D, Barber GP, Casper J, Clawson H, Cline MS, Diekhans M, Dreszer TR, Fujita PA, Guruvadoo L, Haeussler M *et al.* The UCSC Genome Browser database: 2014 update. *Nucleic Acids Res* 2014;**42**:D764–D770.

Katayama S, Tohonen V, Linnarsson S, Kere J. SAMstrt: statistical test for differential expression in single-cell transcriptome with spike-in normalization. *Bioinformatics* 2013;**29**:2943–2945.

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

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W *et al*. Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860–921.

Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;**10**:R25.

Li J, Tibshirani R. Finding consistent patterns: a nonparametric approach for identifying differential expression in RNA-Seq data. *Stat Methods Med Res* 2013;**22**:519–536.

Medina I, Carbonell J, Pulido L, Madeira SC, Goetz S, Conesa A, Tarraga J, Pascual-Montano A, Nogales-Cadenas R, Santoyo J *et al*. Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling. *Nucleic Acids Res* 2010; **38**:W210–W213.

Ploner A. Heatplus: Heatmaps with row and/or column covariates and colored clusters. R package version 2.14.0. 2014.

Pruitt KD, Tatusova T, Maglott DR. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* 2005;**33**:D501–D504.