

The American Journal of Human Genetics, Volume 96

Supplemental Data

Biased Allelic Expression in Human Primary Fibroblast Single Cells

**Christelle Borel, Pedro G. Ferreira, Federico Santoni, Olivier Delaneau, Alexandre Fort,
Konstantin Y. Popadin, Marco Garieri, Emilie Falconnet, Pascale Ribaux, Michel
Guipponi, Ismael Padioleau, Piero Carninci, Emmanouil T. Dermitzakis, and Stylianos E.
Antonarakis**

Figure S1.

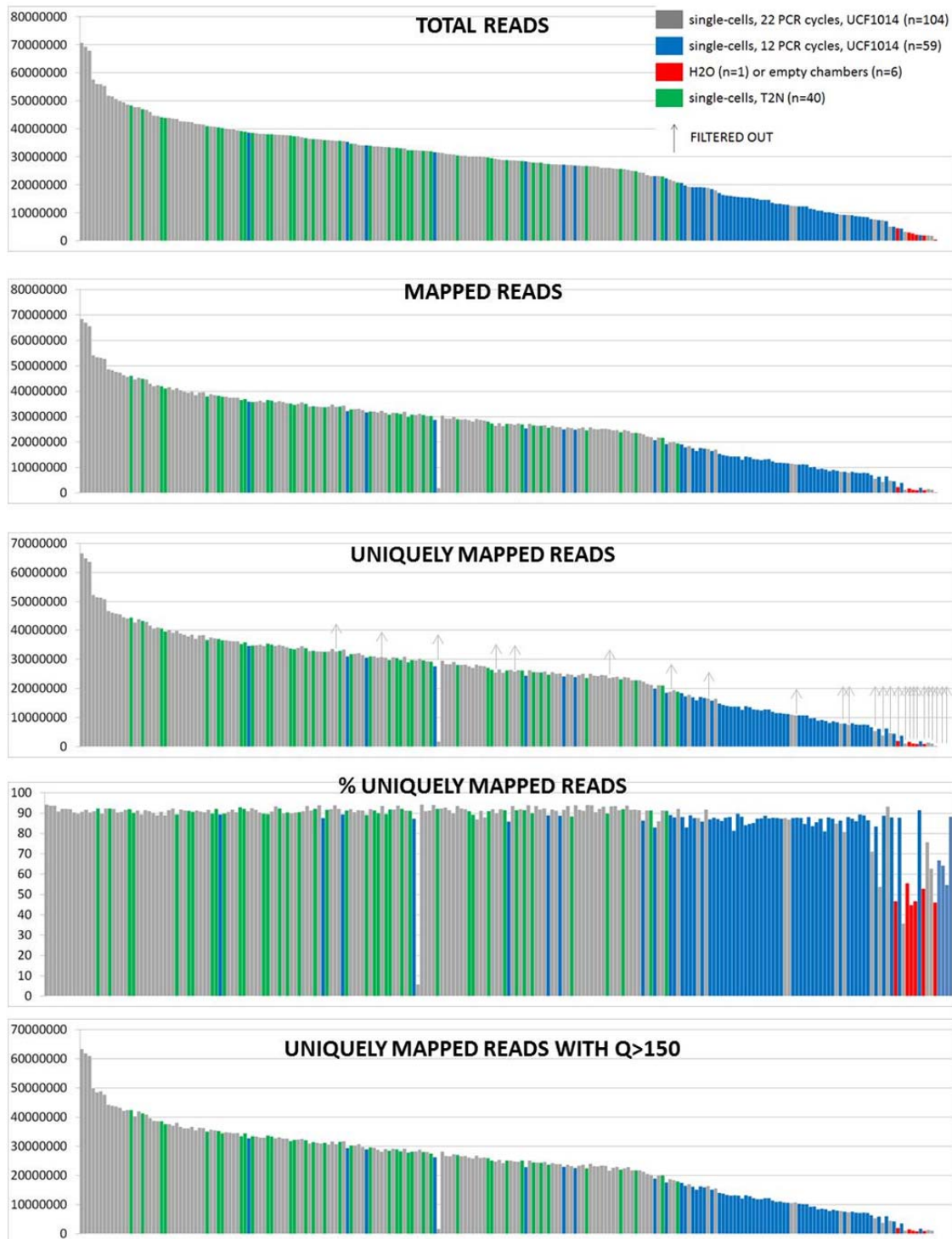


Figure S1. Read distribution for single-cell samples (UCF1014 and T2N)

Histograms displayed read distribution per sample (y-axis, number of reads or percentage of reads). We used color code to differentiate single-cell samples amplified with 12 PCR cycles (blue) or with 22 PCR cycles (grey, standard protocol). We included one sample containing water instead of pre-amplified cDNA and samples from empty capture chamber, i.e. chambers containing no cell after single-cell capture procedure (red). Samples colored in green are single-cell samples from T2N fibroblasts, whereas all the other single-cell samples have been prepared from UCF1014 fibroblasts.

Figure S2.

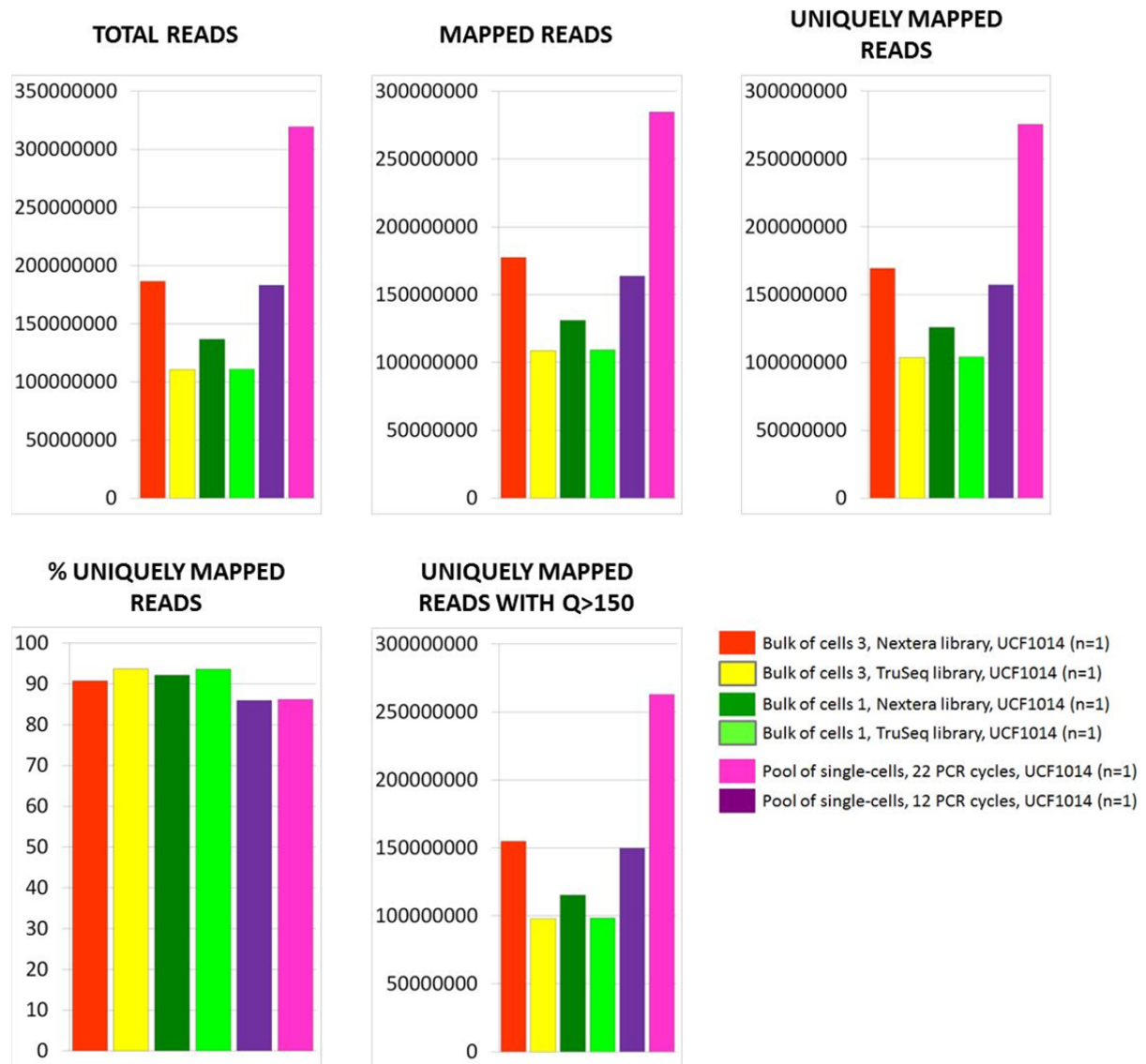


Figure S2. Read distribution for bulk of cells and pool of single-cells (UCF1014). Histograms displayed read distribution per sample (y-axis, number of reads or percentage of reads). We used color code to differentiate samples. See Figure 1 and method section summarizing the experimental workflow and describing the different types of samples.

Figure S3.

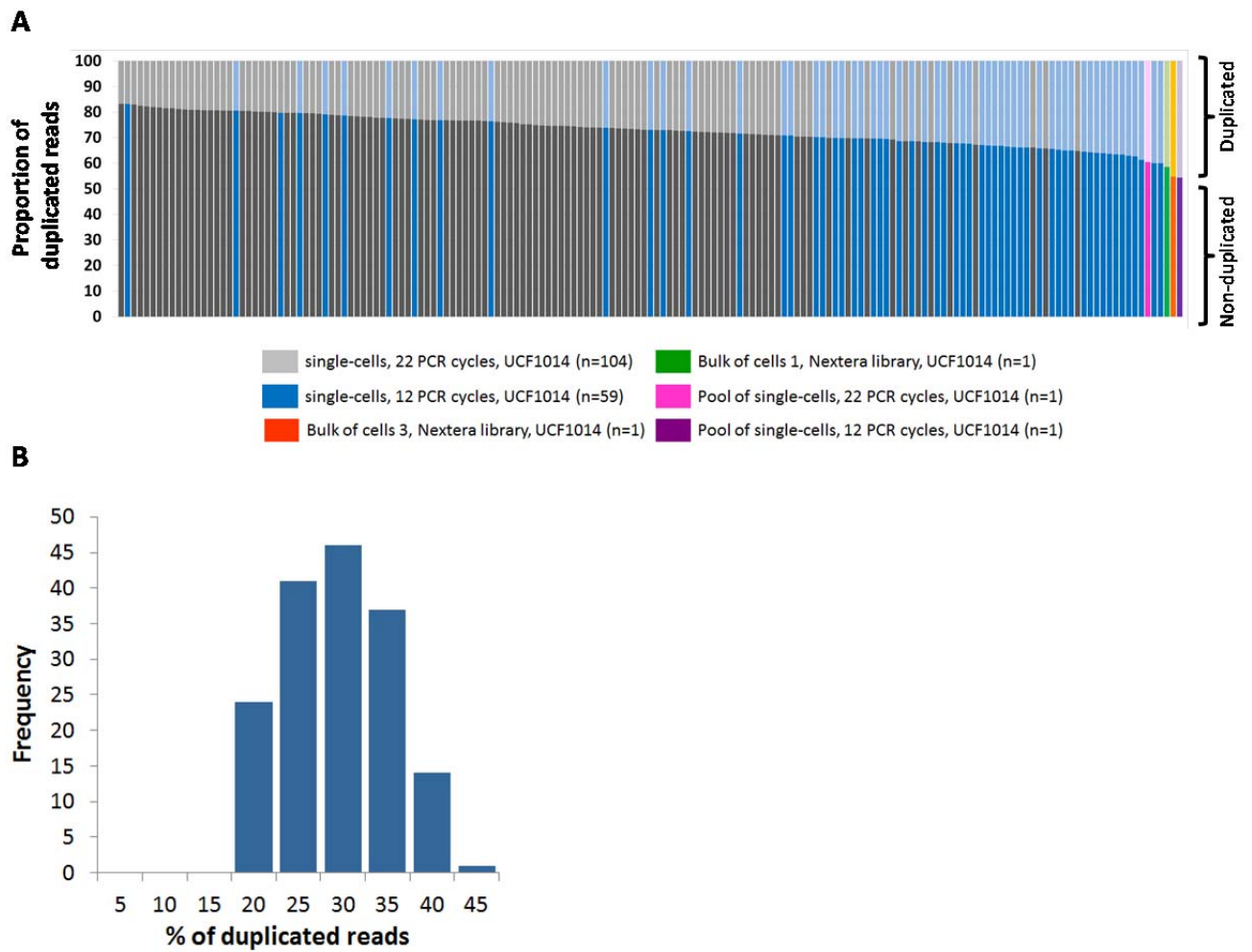


Figure S3. Read duplicates (UCF1014).

(a) Read duplicate refers to reads with identical coordinates. We determined read duplicates for all the samples and we depicted their proportion. Single-cell samples have an average of 73.2% of non-duplicated reads and standard deviation of 5.6%.

(b) Frequency graph shows the distribution of the % of read duplicates across the single-cell samples.

Figure S4.

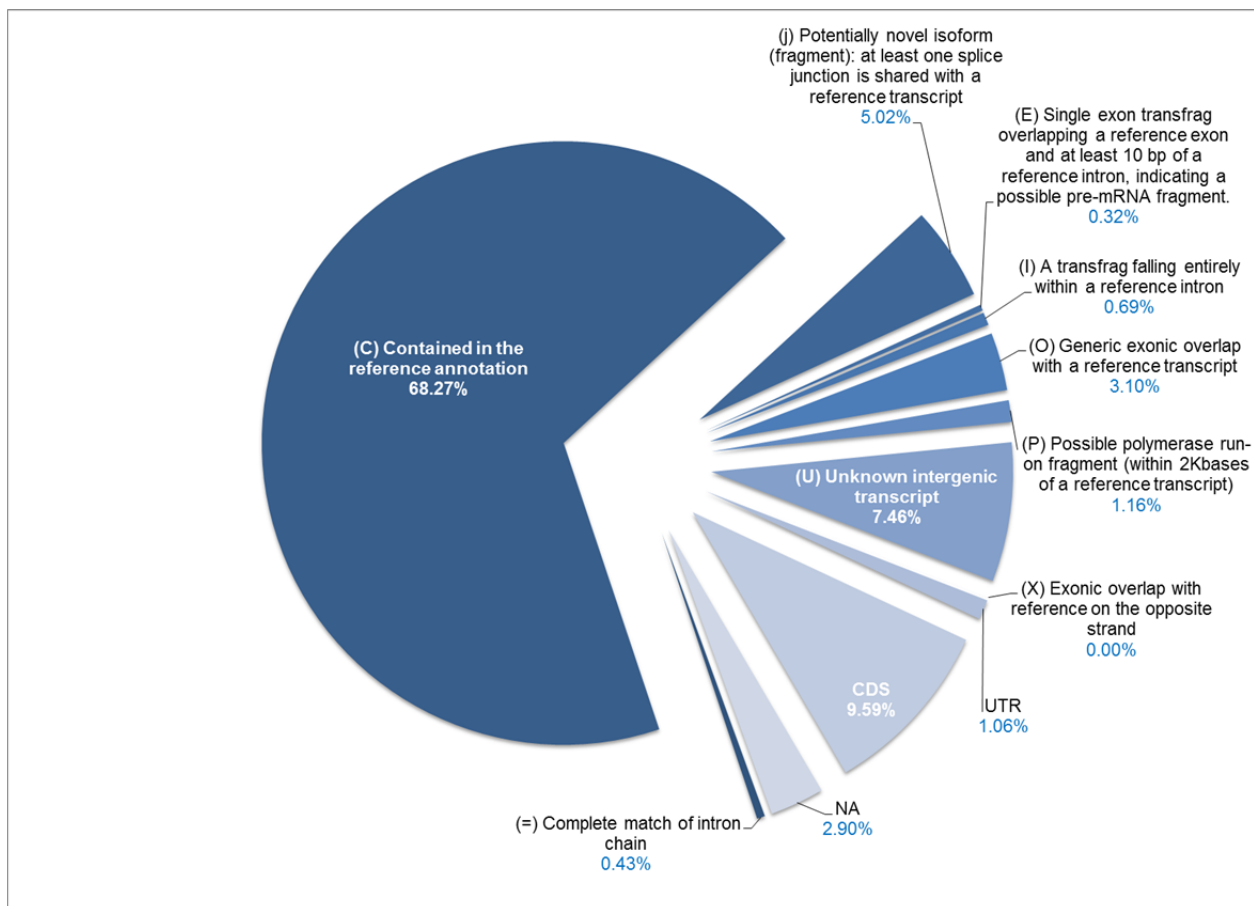


Figure S4. Genomic annotation of 83576 hetSNVs (UCF1014) (>16 reads at position) after cufflinks novel assembly. The pie chart shows the percentage of hetSNVs classified in each Cufflinks class code described by Cuffcompare.

Figure S5.

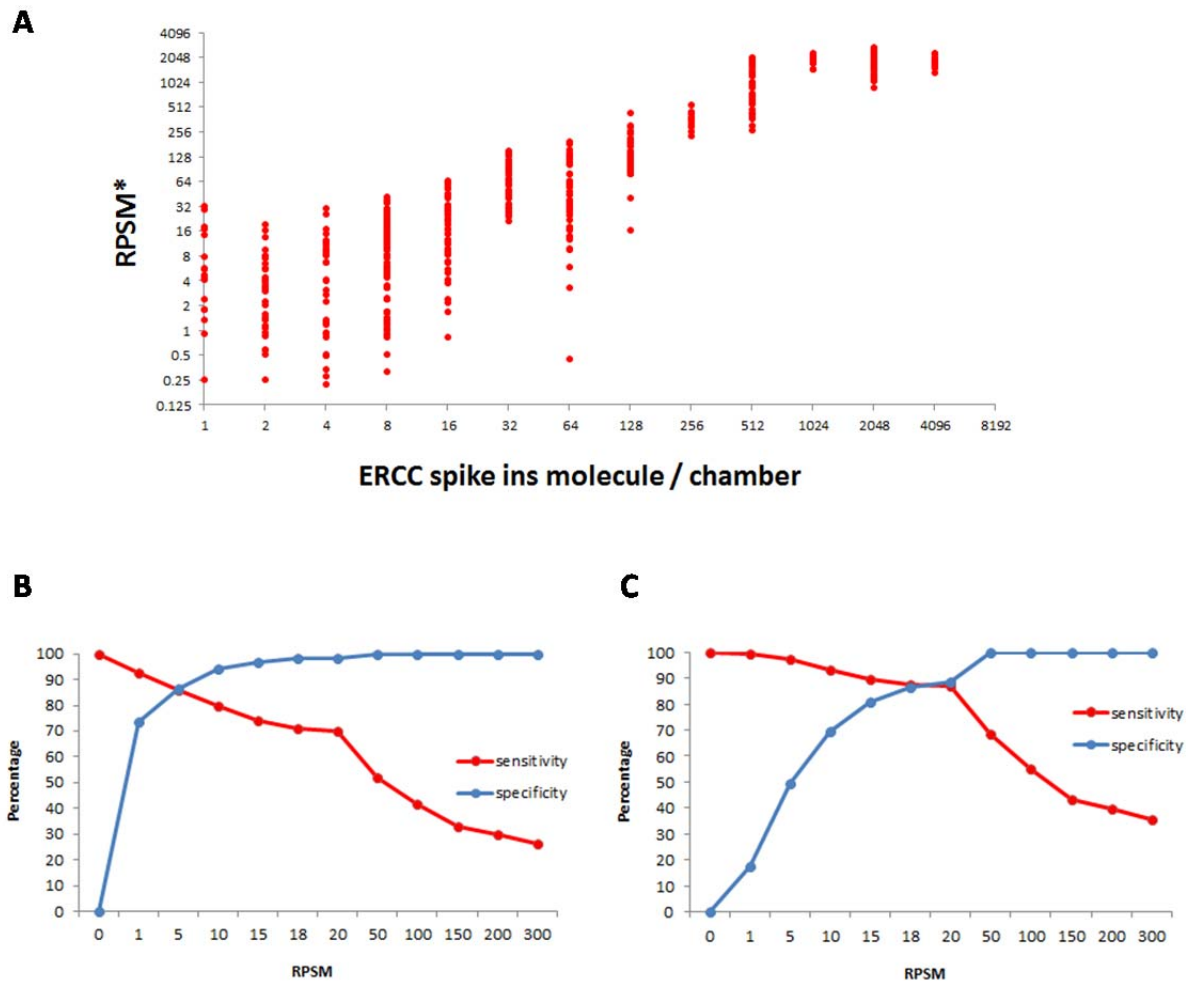


Figure S5. Spike-in experiment

During a single-cell capture experiment, we added to the lysis mix 92 ERCC synthetic spike-in controls at known concentrations. The lysis mix is then uniformly distributed in every chambers. We sequenced 12 spiked fibroblasts from 12 different chambers. (a) The plot shows the number of predicted ERCC spike-in molecule number (see methods) versus the observed normalized read number expressed in RPSM units for a nucleotide located in the middle of the spike sequence (* RPSM, reads at a single nucleotide position per sequencing read length (in kb) and per million mapped reads). Additionally, we calculated the sensitivity and specificity for 4 spike-in molecules (b) and for 8 spike-in molecules (c) at different RPSM levels.

Figure S6.

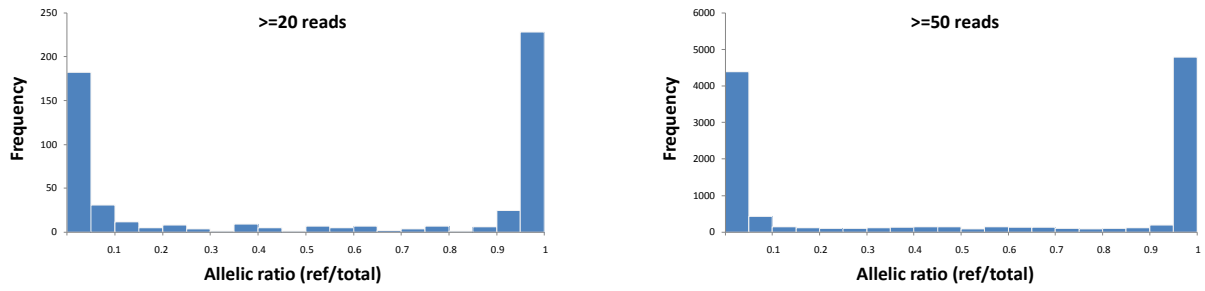


Figure S6. Allelic ratio and read coverage at hetSNV position (UCF1014).

Frequency distribution histograms of the allelic ratio (reference reads / total reads) according to the read coverage at hetSNVs position (RPSM>20, 163 UCF1014 single cells).

Figure S7.

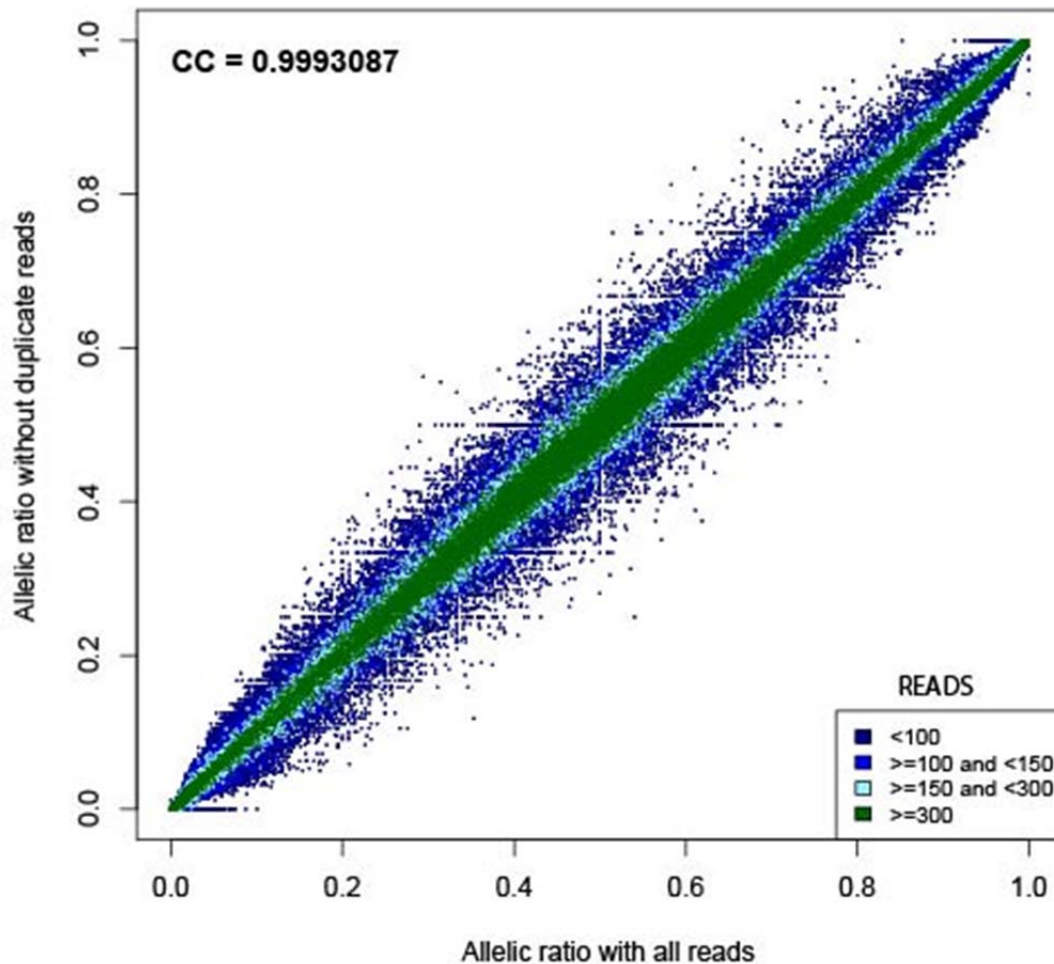
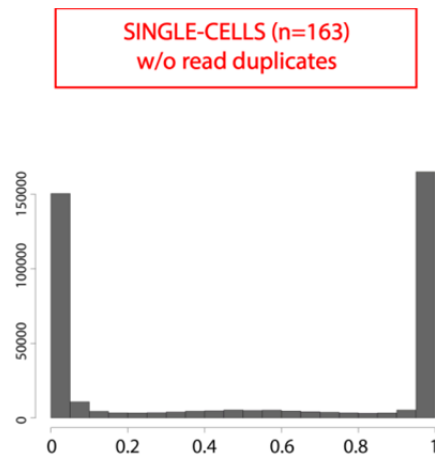


Figure S7. Effect of duplicates reads on allelic ratio.

Pairwise comparison of the allelic ratio after removing the duplicate reads (Y-axis) and all the reads (X-axis). Pearson coefficient of correlation is indicated at the top of the scatter plot. Points are color coded according to the read coverage at hetSNVs position. (RPSM threshold is not applied; 163 UCF1014 single cells).

Figure S8.

A



B

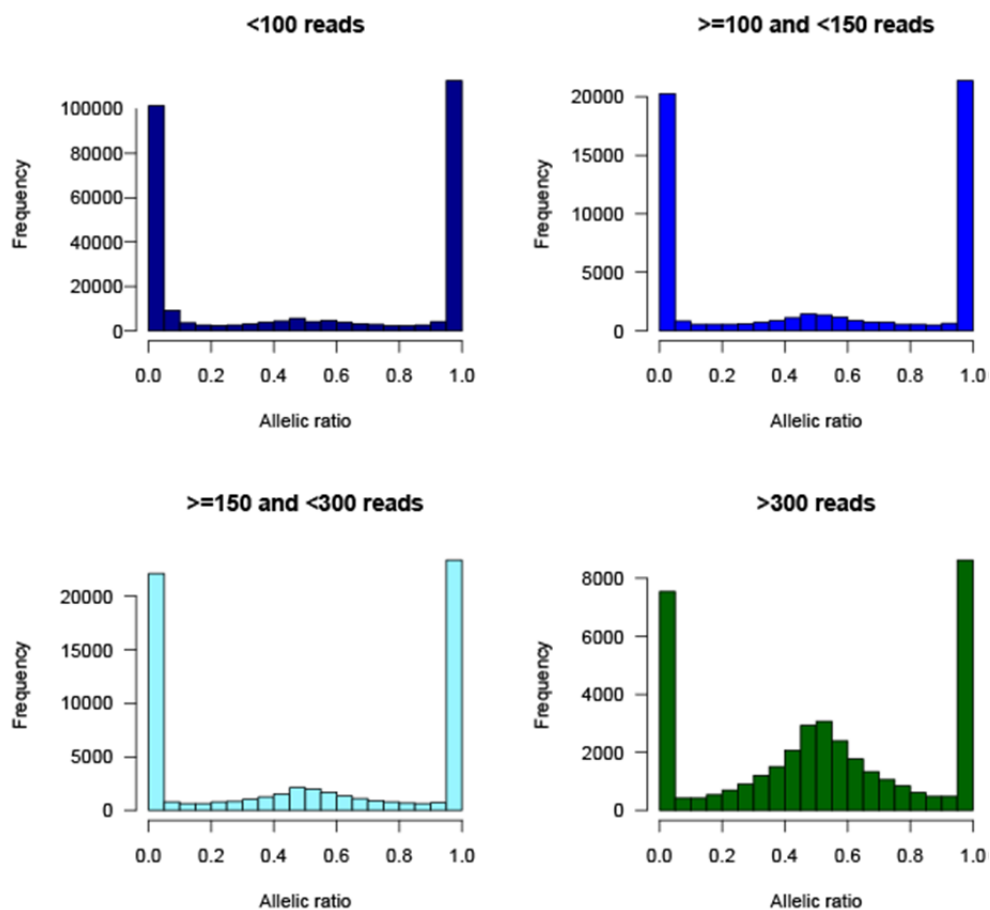


Figure S8. Allelic ratio and read coverage at hetSNV position (UCF1014) after removing duplicate reads. (a) Frequency distribution histogram of the allelic ratio of single cells (reference reads/ total reads) after removing duplicate reads. **(b)** Frequency distribution according to the read coverage. (RPSM threshold is not applied; 163 UCF1014 single cells).

Figure S9.

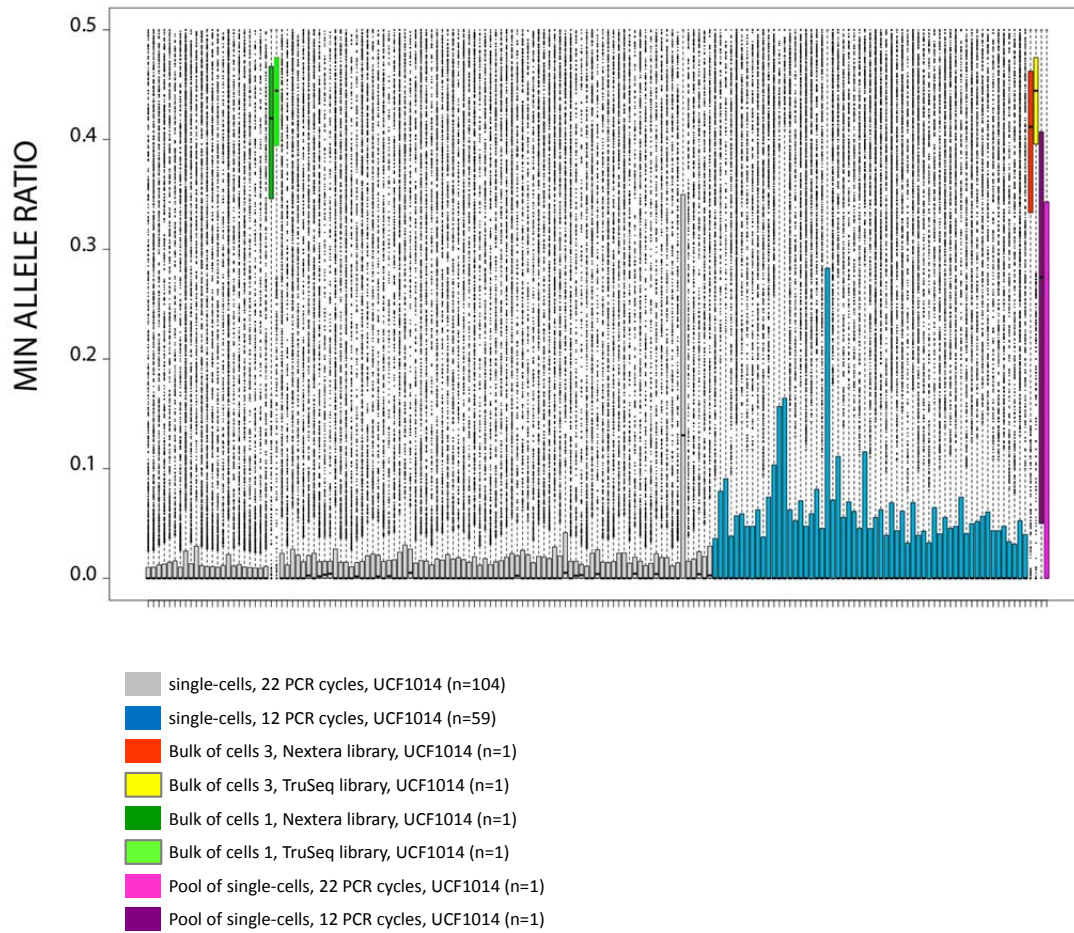


Figure S9. Distribution of minimum allelic ratio across the samples. The minimum allelic ratio is the absolute value of the difference between 0.5 and the allelic ratio (reference reads / total reads). Graph is made of side-by-side boxplots (x-axis, samples). RPSM threshold is not applied.

Figure S10.

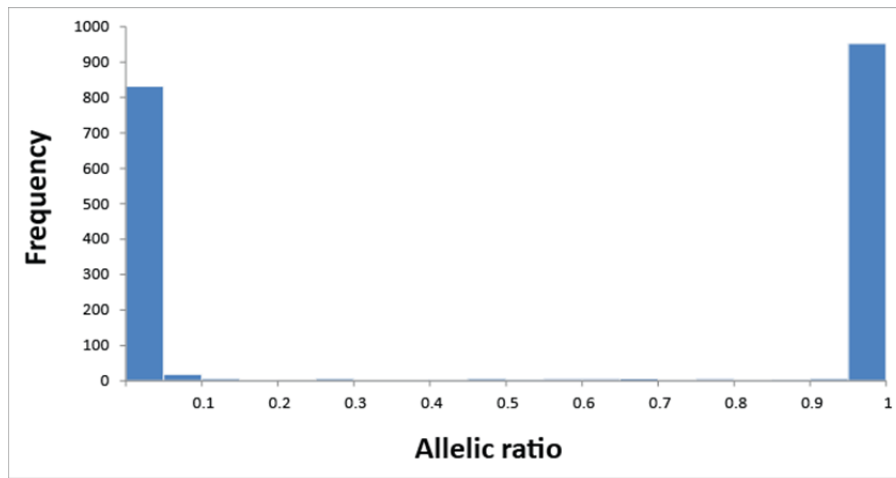
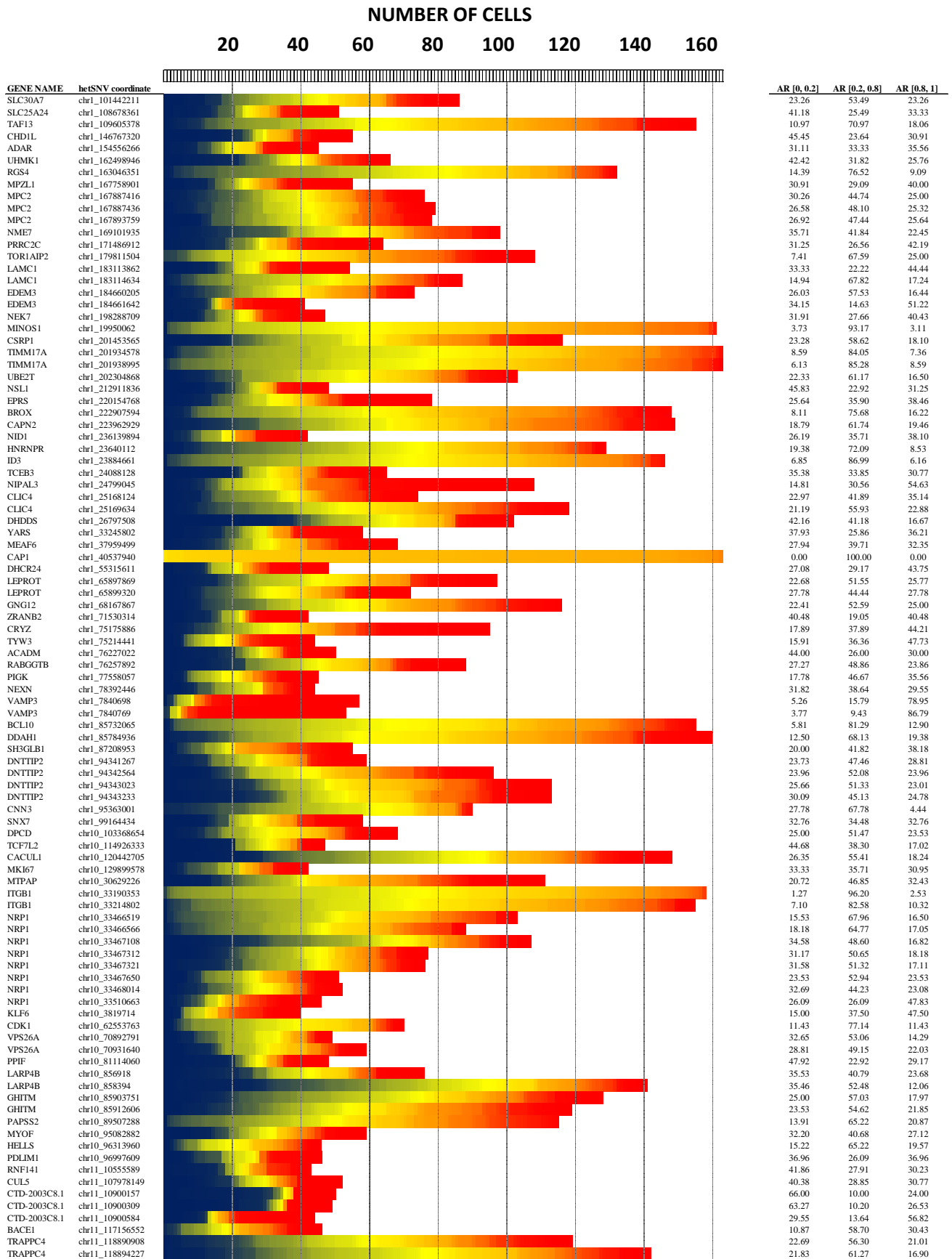
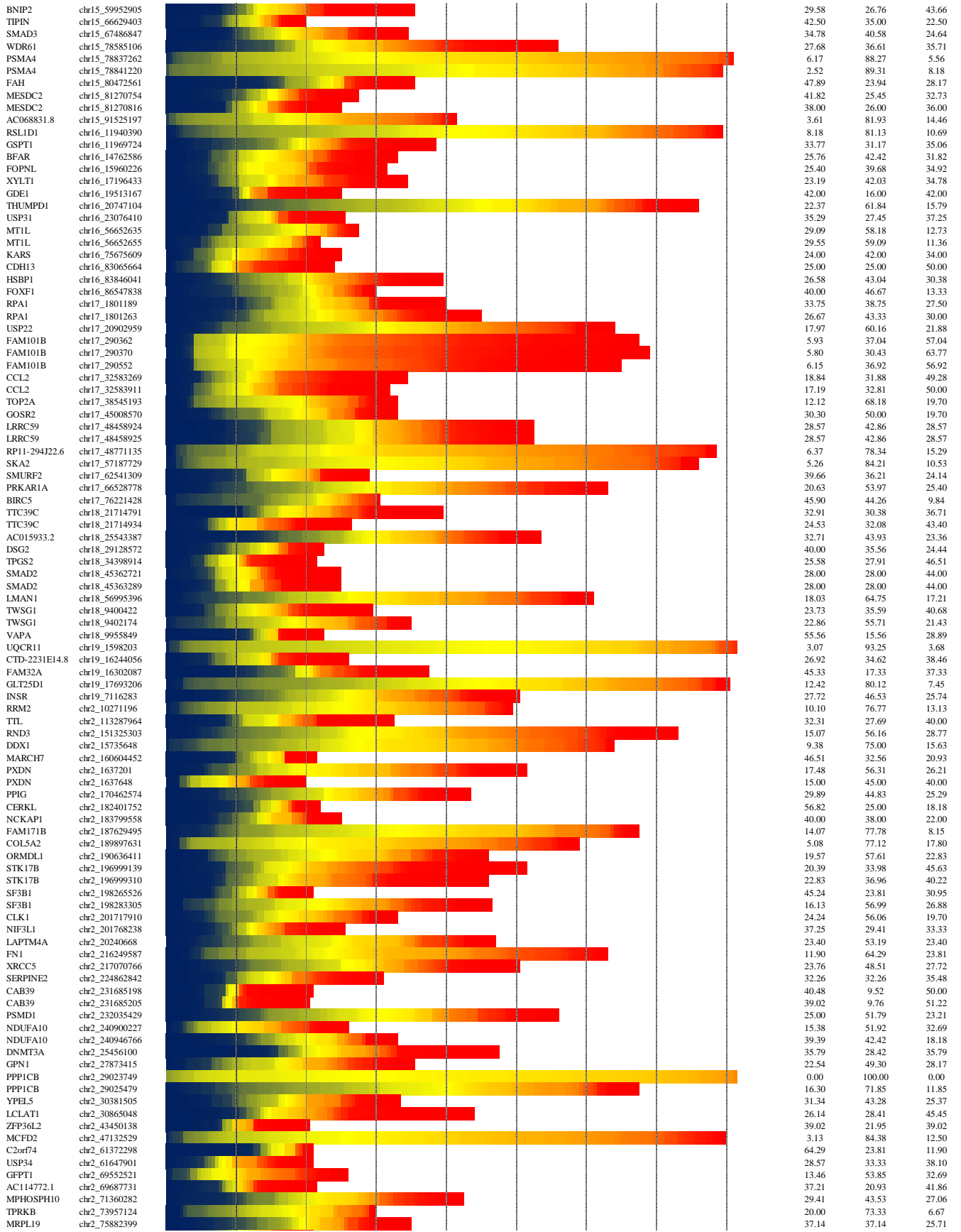


Figure S10. Allelic ratio in single cells on chromosome X (UCF1014). Frequency distribution histograms of the allelic ratio (reference reads / total reads) for hetSNVs located on chromosome X (RPSM>20, 163 UCF1014 single cells). As expected, due to X-inactivation, the transcription is random (one allele or the other with no specification of the parental origin) and monoallelic across all the single cells.

Figure S11.

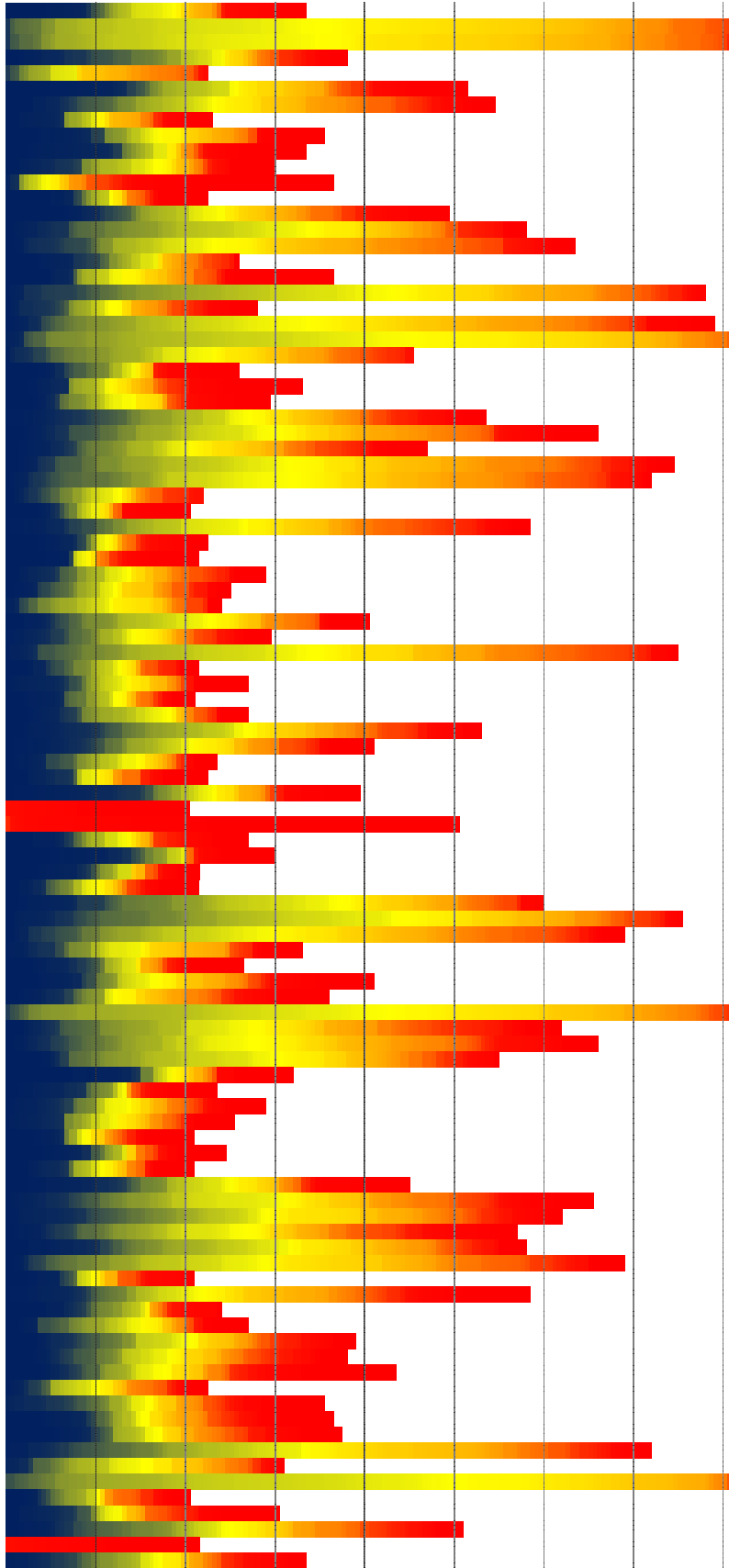


DPAGT1	chr11_118967291									34.00	38.00	28.00
DPAGT1	chr11_118967758									30.36	35.71	33.93
USP47	chr11_11976628									33.90	35.59	30.51
RP11-166D19.1	chr11_121960257									55.83	43.56	0.61
CLMP	chr11_122943495									34.88	13.95	51.16
CLMP	chr11_122943802									24.32	31.08	44.59
SRFR	chr11_126133208									24.29	31.43	44.29
CD44	chr11_35252926									3.09	93.83	3.09
CD44	chr11_35253272									3.09	96.91	0.00
RRM1	chr11_4141132									27.63	44.74	27.63
ATL3	chr11_63393243									4.17	29.17	66.67
ATL3	chr11_63395639									2.91	24.27	72.82
NEAT1	chr11_65205958									32.56	39.53	27.91
CCND1	chr11_69466737									36.07	32.79	31.15
RAB6A	chr11_73388363									36.59	24.39	39.02
MRPL48	chr11_73575499									41.56	35.06	23.38
MRPL48	chr11_73575504									42.86	33.77	23.38
PRCP	chr11_82535461									36.07	40.98	22.95
EED	chr11_85968623									47.62	21.43	30.95
IPO7	chr11_9468367									17.57	59.46	22.97
RAD52	chr12_1021557									73.81	2.38	23.81
NTSDC3	chr12_104167185									28.00	34.00	38.00
NTSDC3	chr12_104167404									27.08	33.33	39.58
NTSDC3	chr12_104168087									24.44	35.56	40.00
C12orf75	chr12_105764984									4.20	61.34	34.45
C12orf23	chr12_107367224									11.95	81.76	6.29
C12orf23	chr12_107367225									11.95	81.76	6.29
C12orf23	chr12_107367439									7.50	83.13	9.38
MAGOHB	chr12_10758703									22.22	64.81	12.96
MAGOHB	chr12_10763236									18.07	66.27	15.66
CSDA	chr12_10851772									18.26	68.70	13.04
WSB2	chr12_118471760									21.74	26.09	52.17
TRIAP1	chr12_120882112									22.55	43.14	34.31
TRIAP1	chr12_120882217									23.08	43.27	33.65
POP5	chr12_121017171									21.74	51.30	26.96
GPRC5A	chr12_13069410									34.62	38.46	26.92
GPRC5A	chr12_13070752									34.09	47.73	18.18
RAN	chr12_131361241									30.00	40.00	30.00
EMPI	chr12_13369574									25.81	38.71	35.48
MGST1	chr12_16516994									20.00	52.14	27.86
GOLT1B	chr12_21669996									19.18	50.68	30.14
GOLT1B	chr12_21670020									31.82	51.52	16.67
MRPS35	chr12_27867727									31.91	38.30	29.79
MRPS35	chr12_27908603									20.00	45.00	35.00
C12orf5	chr12_4462040									28.81	37.29	33.90
ARID2	chr12_46300100									22.14	58.57	19.29
AMIGO2	chr12_47471439									32.00	26.00	42.00
TMEM106C	chr12_48359984									39.08	36.78	24.14
TMEM106C	chr12_48362412									35.37	42.68	21.95
LIMA1	chr12_50570127									30.89	47.97	21.14
LIMA1	chr12_50570519									20.45	42.05	37.50
LIMA1	chr12_50616346									13.77	78.26	7.97
CBX5	chr12_54625518									25.96	47.12	26.92
NACA	chr12_57108203									1.84	96.93	1.23
USP15	chr12_62696599									19.35	50.54	30.11
HMG2	chr12_66358347									35.29	27.45	37.25
HMG2	chr12_66359752									24.36	37.18	38.46
COPS7A	chr12_6840371									28.26	36.96	34.78
KRR1	chr12_75893486									0.00	7.58	92.42
KRR1	chr12_75900588									26.32	43.86	29.82
C12orf29	chr12_88440676									25.45	23.64	50.91
DCN	chr12_91539687									27.66	46.81	25.53
EEA1	chr12_93166538									39.33	32.58	28.09
EEA1	chr12_93167435									24.62	40.00	35.38
MRPL42	chr12_93862989									19.18	67.12	13.70
TMPO	chr12_98943585									12.87	68.32	18.81
CDC16	chr13_115030714									54.17	12.50	33.33
EXOSC8	chr13_37580139									10.87	70.29	18.84
GTF2F2	chr13_45841502									47.27	27.27	25.45
RBM26	chr13_79894736									24.69	50.62	24.69
STK24	chr13_99104689									40.43	40.43	19.15
PPP2R5C	chr14_102323093									19.33	57.14	23.53
ZFYVE21	chr14_104199580									45.24	26.19	28.57
DAD1	chr14_23033838									0.61	96.93	2.45
C14orf19	chr14_23567369									19.44	57.64	22.92
PSME1	chr14_24606928									25.00	53.00	22.00
SEC23A	chr14_39517925									36.99	38.36	24.66
FRMD6	chr14_52196115									24.77	50.46	24.77
CNIH	chr14_54894296									18.05	69.92	12.03
DLGAP5	chr14_55655692									16.18	69.12	14.71
KTN1	chr14_56078739									17.31	69.23	13.46
JKAMP	chr14_59971039									16.20	59.15	24.65
DHRS7	chr14_60611648									18.84	53.62	27.54
DHRS7	chr14_60611675									16.46	55.70	27.85
ELMSAN1	chr14_74182810									13.53	55.64	30.83
EIF2B2	chr14_75476071									57.78	20.00	22.22
GPATCH2L	chr14_76645514									37.78	17.78	44.44
RP11-66B24.4	chr15_101455840									40.00	44.71	15.29
CHSY1	chr15_101716583									36.36	44.32	19.32
CHSY1	chr15_101717065									38.27	39.51	22.22
NDUFAF1	chr15_41679685									40.00	20.00	40.00
TMEM87A	chr15_42529665									46.51	23.26	30.23
FBN1	chr15_48702457									13.79	63.45	22.76
FBN1	chr15_48702873									20.45	51.52	28.03
USP8	chr15_50769520									39.58	25.00	35.42
ARPP19	chr15_52843442									38.30	38.30	23.40
BNIP2	chr15_59951573									43.75	26.25	30.00
BNIP2	chr15_59951997									41.33	29.33	29.33
BNIP2	chr15_59952253									47.44	25.64	26.92
BNIP2	chr15_59952341									48.15	24.69	27.16



MRPL19	chr2_75882576									35.71	35.71	28.57
CHMP3	chr2_86731290									38.10	16.67	45.24
RNF103-CHMP3	chr2_86756381									35.38	29.23	35.38
MRPS5	chr2_95775761									29.41	41.18	29.41
STARD7	chr2_96850885									14.19	67.74	18.06
FKBP1A	chr20_1350672									9.87	77.63	12.50
ITCH	chr20_33095891									25.00	30.00	45.00
PROCR	chr20_33764632									29.51	40.98	29.51
RP4-564F22.2	chr20_37059749									8.00	14.00	78.00
SRSF6	chr20_42091357									3.07	91.41	5.52
YWHAB	chr20_43530234									8.92	74.52	16.56
CSEIL	chr20_47685320									19.48	50.65	29.87
ZNFY1-AS1	chr20_47895702									5.92	71.71	22.37
ADNP	chr20_49506195									16.23	69.48	14.29
AURKA	chr20_54945211									30.43	47.83	21.74
RPS21	chr20_60962943									0.61	98.77	0.61
PCMTD2	chr20_62906433									27.45	33.33	39.22
PCMTD2	chr20_62906514									31.37	29.41	39.22
MRPL39	chr21_26969703									30.26	53.95	15.79
MRPL39	chr21_26978950									25.71	58.57	15.71
IFNAR1	chr21_34715699									31.71	26.83	41.46
SON	chr21_34926260									34.55	23.64	41.82
WRB	chr21_40769290									11.63	62.79	25.58
MAPK1	chr22_22115004									43.14	27.45	29.41
C22orf28	chr22_32783904									19.13	53.04	27.83
C22orf28	chr22_32795641									28.57	34.29	37.14
TIMP3	chr22_33253280									18.58	55.75	25.66
TIMP3	chr22_33253292									18.64	56.78	24.58
NHP2L1	chr22_42070793									32.61	30.43	36.96
ARFGAP3	chr22_43218397									16.36	52.73	30.91
TFG	chr3_100467018									25.74	43.56	30.69
TFG	chr3_100467673									20.17	52.10	27.73
TFG	chr3_100467676									19.33	52.94	27.73
CD200	chr3_112059768									31.82	40.91	27.27
CD200	chr3_112063850									25.00	44.74	30.26
CCDC80	chr3_112359986									2.56	65.38	32.05
B4GALT4	chr3_118931389									40.74	18.52	40.74
POGLUT1	chr3_119213220									24.35	45.22	30.43
TIMMDC1	chr3_119242803									17.65	64.71	17.65
FSTL1	chr3_120114859									17.43	63.30	19.27
KPNA1	chr3_122140919									22.50	22.50	55.00
UMPS	chr3_124462808									30.51	37.29	32.20
TMEM43	chr3_14184719									43.10	27.59	29.31
GYGI	chr3_148745236									27.03	43.24	29.73
HLTF	chr3_148748380									39.13	19.57	41.30
FGD5-AS1	chr3_14984617									26.79	50.00	23.21
SERP1	chr3_150260853									41.67	29.17	29.17
SELT	chr3_150345690									5.56	72.22	22.22
EAF1	chr3_15483830									27.42	46.77	25.81
SSR3	chr3_156259966									9.49	82.91	7.59
SSR3	chr3_156266774									1.23	93.25	5.52
LXN	chr3_158384463									22.41	51.72	25.86
TRIM59	chr3_160152299									25.33	44.00	30.67
TRIM59	chr3_160156951									26.00	48.00	26.00
DPH3	chr3_16301564									28.57	16.67	54.76
DPH3	chr3_16305674									13.89	51.39	34.72
AC008040.1	chr3_169806897									39.66	17.24	43.10
SKIL	chr3_170112307									6.94	79.86	13.19
SKIL	chr3_170113481									10.32	72.90	16.77
SKIL	chr3_170113533									11.11	70.59	18.30
SKIL	chr3_170113601									12.75	67.11	20.13
DCUN1D1	chr3_182659953									34.21	35.53	30.26
TFRC	chr3_195776454									20.28	56.64	23.08
RPL15	chr3_23961353									12.14	67.14	20.71
RPL15	chr3_23961621									14.29	70.13	15.58
CDCP1	chr3_45125067									48.08	26.92	25.00
FYCO1	chr3_45959759									34.04	29.79	36.17
MAP4	chr3_47958037									42.86	35.71	21.43
FAM208A	chr3_56654703									57.69	15.38	26.92
ARF4	chr3_57557620									3.07	89.57	7.36
PDHB	chr3_58413662									22.68	51.55	25.77
PDHB	chr3_58413669									26.04	52.08	21.88
THUMPD3	chr3_9425911									31.91	19.15	48.94
THUMPD3	chr3_9425944									48.89	17.78	33.33
THUMPD3	chr3_9426224									47.06	21.57	31.37
DCBLD2	chr3_98516386									28.77	49.32	21.92
ZNF518B	chr4_10442700									26.67	36.00	37.33
PPA2	chr4_106317429									35.29	38.24	26.47
CAMK2D	chr4_114375133									33.66	51.49	14.85
C4orf3	chr4_120219534									33.33	40.00	26.67
MAD2L1	chr4_120981214									20.62	70.10	9.28
NAA15	chr4_140310520									26.06	66.20	7.75
FAM198B	chr4_159045843									22.22	62.22	15.56
PPID	chr4_159630817									30.95	23.81	45.24
MSMO1	chr4_166264119									9.23	81.54	9.23
ANXA10	chr4_169083694									37.50	22.50	40.00
SLBP	chr4_1694809									25.00	38.46	36.54
PALLD	chr4_169849389									19.55	58.65	21.80
SPCS3	chr4_177249681									9.42	60.87	29.71
SPCS3	chr4_177249913									39.72	55.32	4.96
DCTD	chr4_183811670									24.62	36.92	38.46
DCTD	chr4_183815688									36.21	34.48	29.31
CASP3	chr4_185548951									47.83	30.43	21.74
KIAA1430	chr4_186111639									33.33	40.74	25.93
UGDH	chr4_39500514									9.33	60.67	30.00
UGDH	chr4_39501086									5.73	65.61	28.66
ATP10D	chr4_47514685									43.48	28.26	28.26
DCUN1D4	chr4_52781759									17.28	38.27	44.44
TMEM165	chr4_56292101									14.29	47.62	38.10
UBA6	chr4_68482401									32.35	38.24	29.41

UBA6 chr4_68482405
 UBA6 chr4_68483472
 UBA6 chr4_68483534
 CXCL6 chr4_74703999
 SEPT11 chr4_77958130
 SEPT11 chr4_77958765
 CCNI chr4_77968968
 PDLIM5 chr4_95588565
 PDLIM5 chr4_95588823
 RAP1GDS1 chr4_99363400
 PJA2 chr5_108714298
 WDR36 chr5_110464687
 ALDH7A1 chr5_125880589
 HSPA9 chr5_137902339
 MATR3 chr5_138665756
 MATR3 chr5_138665792
 NR3C1 chr5_142661595
 YIPF5 chr5_143538589
 YIPF5 chr5_145539142
 LARS chr5_145543972
 CTB-89H12.4 chr5_148873854
 SPARC chr5_151043111
 CLINT1 chr5_157213497
 MYO10 chr5_16666180
 NPM1 chr5_170819887
 KIAA1191 chr5_175773287
 HNRNPAB chr5_177637882
 NUP155 chr5_37288535
 PELO chr5_52097764
 NDUFS4 chr5_52942197
 NDUFS4 chr5_52979097
 SREK1IP1 chr5_64023981
 MAP1B chr5_71492734
 NSA2 chr5_74069863
 WDR41 chr5_76734084
 AP3B1 chr5_77298619
 VCAN chr5_82876271
 VCAN chr5_82877682
 EDIL3 chr5_83237255
 EDIL3 chr5_83238236
 RHOBTB3 chr5_95084131
 GLRX chr5_95149966
 ERAP1 chr5_96110211
 R1OK2 chr5_96513471
 RGM5 chr5_98131505
 DSE chr6_116758932
 GJA1 chr6_121770464
 SERINC1 chr6_122765026
 SERINC1 chr6_122765948
 ECHDC1 chr6_127610021
 SNORD100 chr6_133138010
 BCLAF1 chr6_136590640
 BCLAF1 chr6_136599404
 PERP chr6_138412612
 VTA1 chr6_142541553
 VTA1 chr6_142541769
 FUCA2 chr6_143825104
 PHACTR2 chr6_144151026
 PHACTR2 chr6_144151767
 MRPL18 chr6_160219337
 THBS2 chr6_169616347
 HMGN4 chr6_26546048
 SERPINB9 chr6_2887652
 SERPINB9 chr6_2887771
 SRSF3 chr6_36570366
 GLO1 chr6_38643796
 EC12 chr6_4116262
 MRPS10 chr6_42185564
 UBR2 chr6_42659979
 MAD2L1BP chr6_43608386
 SLC35B2 chr6_44222047
 SLC35B2 chr6_44222109
 ELOVL5 chr6_53133964
 DST chr6_56557411
 DST chr6_56557822
 SSR1 chr6_7285193
 SSR1 chr6_7288297
 SSR1 chr6_7288341
 SSR1 chr6_7288544
 SSR1 chr6_7288613
 SSR1 chr6_7289166
 TTK chr6_80751942
 PGM3 chr6_83880167
 ORC3 chr6_88377049
 UBE2J1 chr6_90037968
 NDUFAF4 chr6_97339078
 NDUFAF4 chr6_97339088
 LAMB1 chr7_107569962
 AC073346.2 chr7_112756994
 TES chr7_115897871
 TES chr7_115897884
 TES chr7_115898177
 CAV1 chr7_116200587
 CAPZA2 chr7_116558716
 CALU chr7_128409580
 CALD1 chr7_134618710
 CREB3L2 chr7_137560697
 MRPS33 chr7_140706157
 TRBC2 chr7_142499071
 PDIA4 chr7_148702362



32.84	37.31	29.85
4.29	91.41	4.29
4.29	90.80	4.91
39.47	38.16	22.37
6.67	84.44	8.89
31.07	40.78	28.16
22.94	53.21	23.85
28.26	43.48	28.26
30.99	45.07	23.94
38.81	23.88	37.31
28.33	45.00	26.67
4.11	20.55	75.34
40.00	26.67	33.33
29.29	44.44	26.26
18.97	65.52	15.52
15.75	63.78	20.47
44.23	38.46	17.31
21.92	35.62	42.47
16.03	74.36	9.62
25.00	39.29	35.71
8.86	74.05	17.09
5.56	93.83	0.62
14.29	70.33	15.38
38.46	25.00	36.54
21.21	30.30	48.48
23.73	37.29	38.98
26.17	49.53	24.30
18.94	61.36	19.70
19.15	52.13	28.72
13.42	70.47	16.11
15.97	66.67	17.36
27.27	45.45	27.27
34.15	26.83	39.02
19.66	52.99	27.35
42.22	22.22	35.56
34.88	18.60	46.51
27.59	43.10	29.31
28.00	46.00	26.00
14.58	68.75	16.67
20.99	64.20	14.81
30.51	40.68	28.81
11.33	71.33	17.33
34.88	37.21	27.91
33.33	38.89	27.78
38.10	38.10	23.81
31.48	46.30	22.22
26.42	51.89	21.70
30.49	43.90	25.61
34.04	46.81	19.15
35.56	31.11	33.33
41.77	32.91	25.32
0.00	0.00	100.00
0.00	0.00	100.00
31.48	29.63	38.89
51.67	18.33	30.00
46.51	25.58	27.91
27.91	34.88	37.21
21.67	68.33	10.00
20.53	68.87	10.60
12.32	69.57	18.12
21.21	54.55	24.24
39.62	26.42	33.96
24.39	41.46	34.15
23.61	40.28	36.11
3.07	92.64	4.29
16.94	54.84	28.23
15.91	62.12	21.97
17.27	67.27	15.45
46.88	25.00	28.13
42.55	17.02	40.43
27.59	41.38	31.03
25.49	47.06	27.45
30.95	28.57	40.48
44.90	18.37	36.73
35.71	35.71	28.57
33.33	40.00	26.67
17.56	58.02	24.43
24.19	58.06	17.74
19.30	48.25	32.46
24.14	59.48	16.38
10.87	68.12	21.01
38.10	21.43	40.48
19.66	48.72	31.62
41.67	31.25	27.08
25.93	48.15	25.93
30.77	42.31	26.92
26.32	43.42	30.26
24.14	32.18	43.68
22.22	51.11	26.67
29.58	32.39	38.03
31.51	30.14	38.36
32.00	33.33	34.67
19.44	64.58	15.97
16.13	70.97	12.90
4.91	93.87	1.23
26.83	36.59	36.59
31.15	19.67	49.18
26.47	50.98	22.55
0.00	0.00	100.00
28.36	41.79	29.85

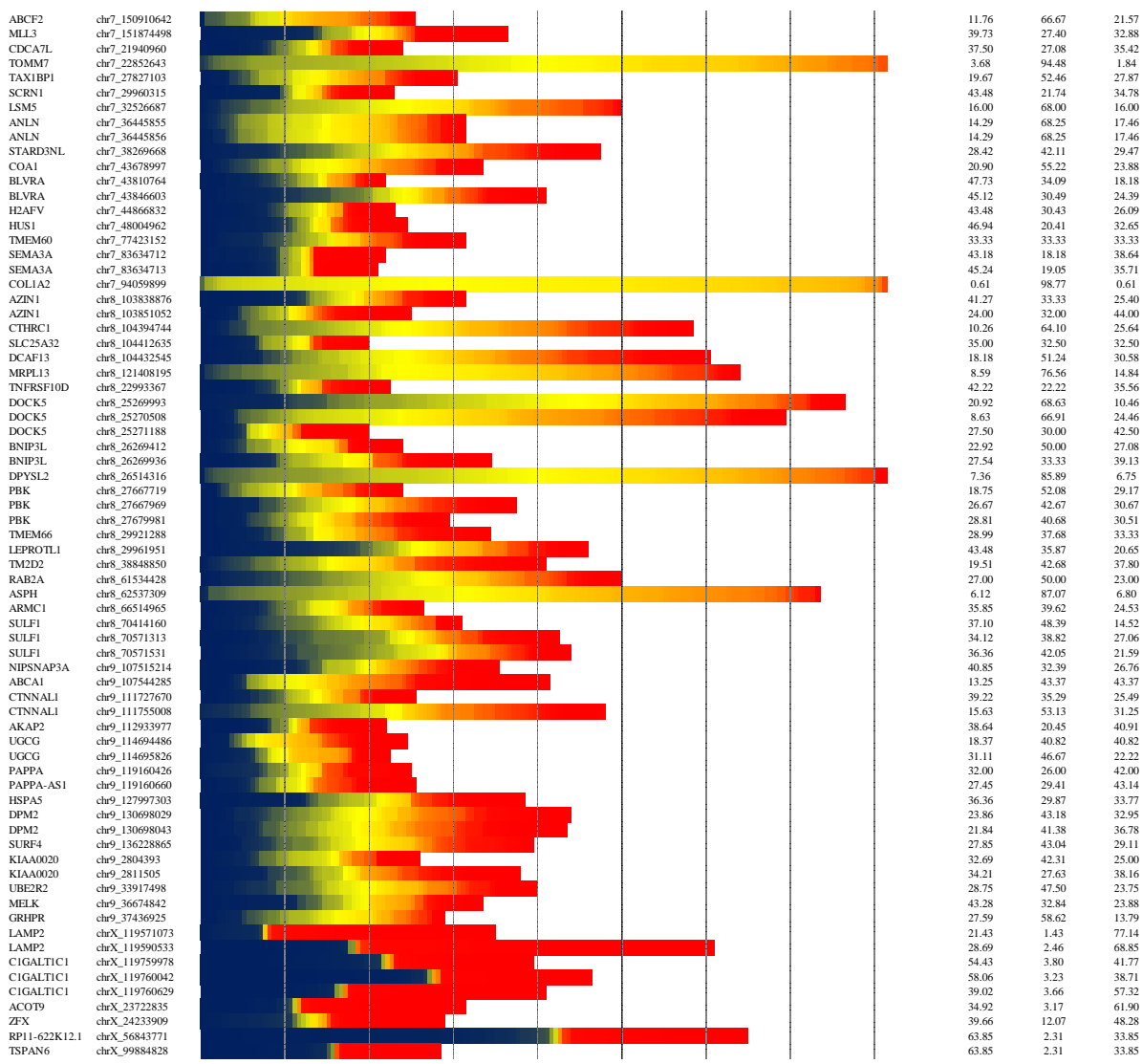


Figure S11. Single-cell allelic ratio distribution per detected genes (UCF1014).

We selected 568 genes for a representative view of the allelic expression in single cells. Index bar indicates the color-coded for the allelic ratio values (reference reads/ total reads). The columns are the gene name, the genomic coordinate of the hetSNV, the color coded column are the single cells, the three last columns are the percentage of cells with an allelic ratio between [0, 0.2], [0.2, 0.8] and [0.8, 1]. We subselected 60 genes out of 568 genes to make the figure 5.