

**JClub**

TG

**Jan 11, 2016**

REPORT


## Single-Cell RNA-Seq Reveals Dynamic, Random Monoallelic Gene Expression in Mammalian Cells


Qiaolin Deng<sup>1,\*</sup>, Daniel Ramsköld<sup>1,2,\*</sup>, Björn Reinius<sup>1,2</sup>, Rickard Sandberg<sup>1,2,†</sup>

 Author Affiliations

<sup>1</sup>Ludwig Institute for Cancer Research, Box 240, 171 77 Stockholm, Sweden.

<sup>2</sup>Department of Cell and Molecular Biology, Karolinska Institutet, 171 77 Stockholm, Sweden.

 <sup>†</sup>Corresponding author. E-mail: [rickard.sandberg@ki.se](mailto:rickard.sandberg@ki.se)

 \* These authors contributed equally to this work.

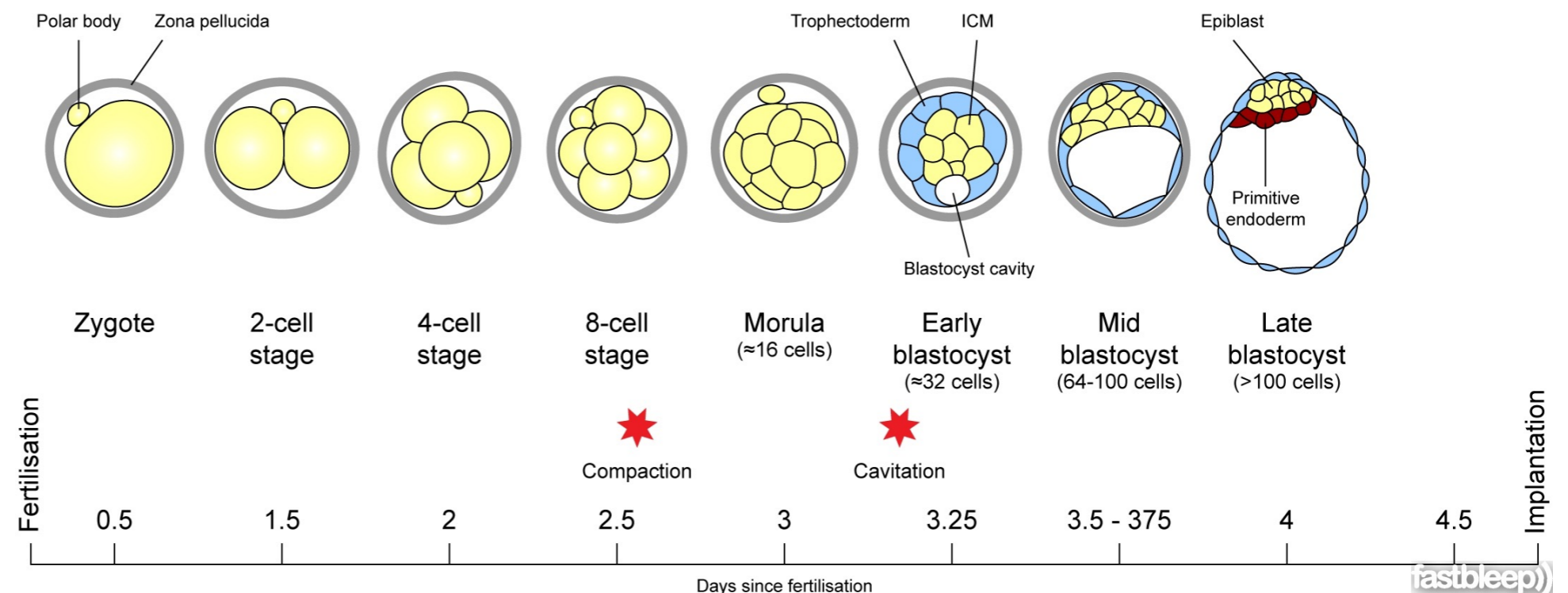
ABSTRACT

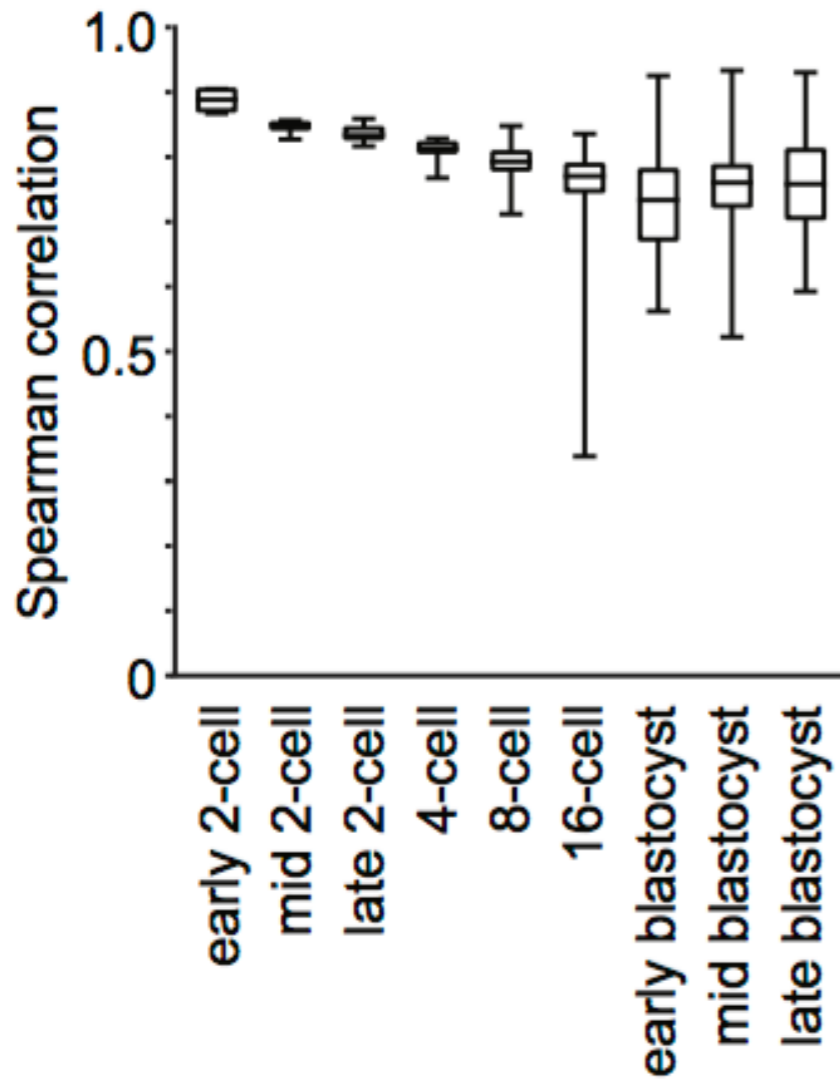
EDITOR'S SUMMARY

Expression from both alleles is generally observed in analyses of diploid cell populations, but studies addressing allelic expression patterns genome-wide in single cells are lacking. Here, we present global analyses of allelic expression across individual cells of mouse preimplantation embryos of mixed background (CAST/EiJ × C57BL/6J). We discovered abundant (12 to 24%) monoallelic expression of autosomal genes and that expression of the two alleles occurs independently. The monoallelic expression appeared random and dynamic because there was considerable variation among closely related embryonic cells. Similar patterns of monoallelic expression were observed in mature cells. Our allelic expression analysis also demonstrates the de novo inactivation of the paternal X chromosome. We conclude that independent and stochastic allelic transcription generates abundant random monoallelic expression in the mammalian cell.

- transcriptome profiles from individual cells with Smart-seq or Smart-seq2 — scRNA-seq
- 269 individual cells dissociated from in vivo F1 embryos (CAST/EiJ x C57BL/6J) — from oocyte to blastocyst stages of mouse preimplantation development
- scRNA-seq on CAST x C57 and C57 x CAST fibroblast cells
- scRNA-seq and bcRNA-seq on liver cells (C57 x CAST)

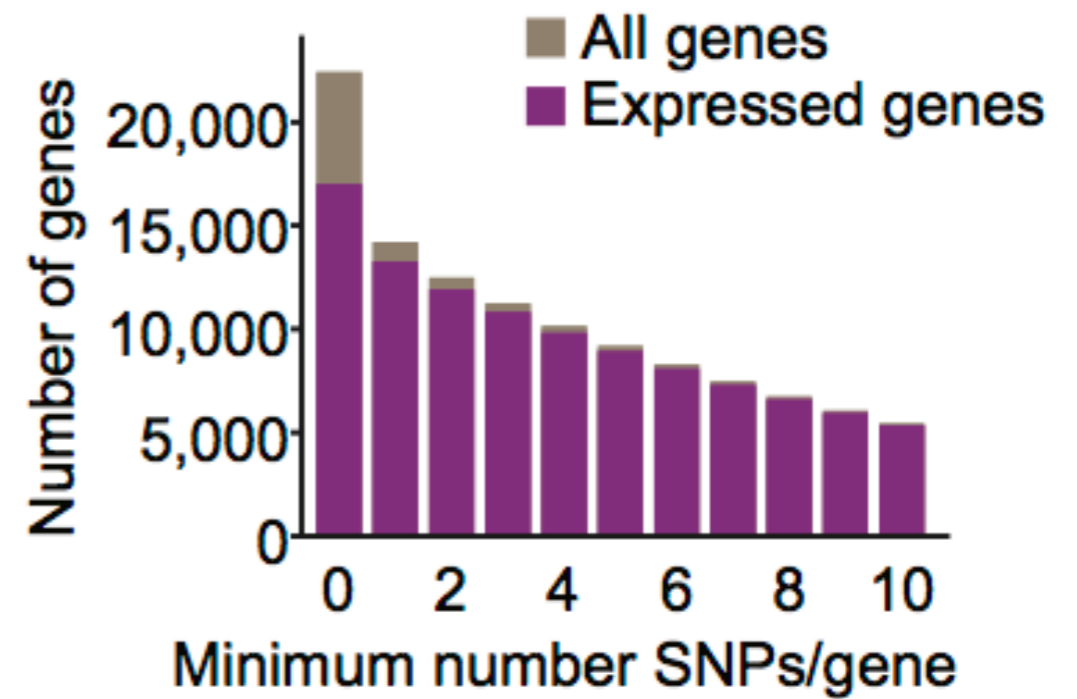
Preimplantation development: initial stages of mammalian development before embryo implants into the mother's uterus

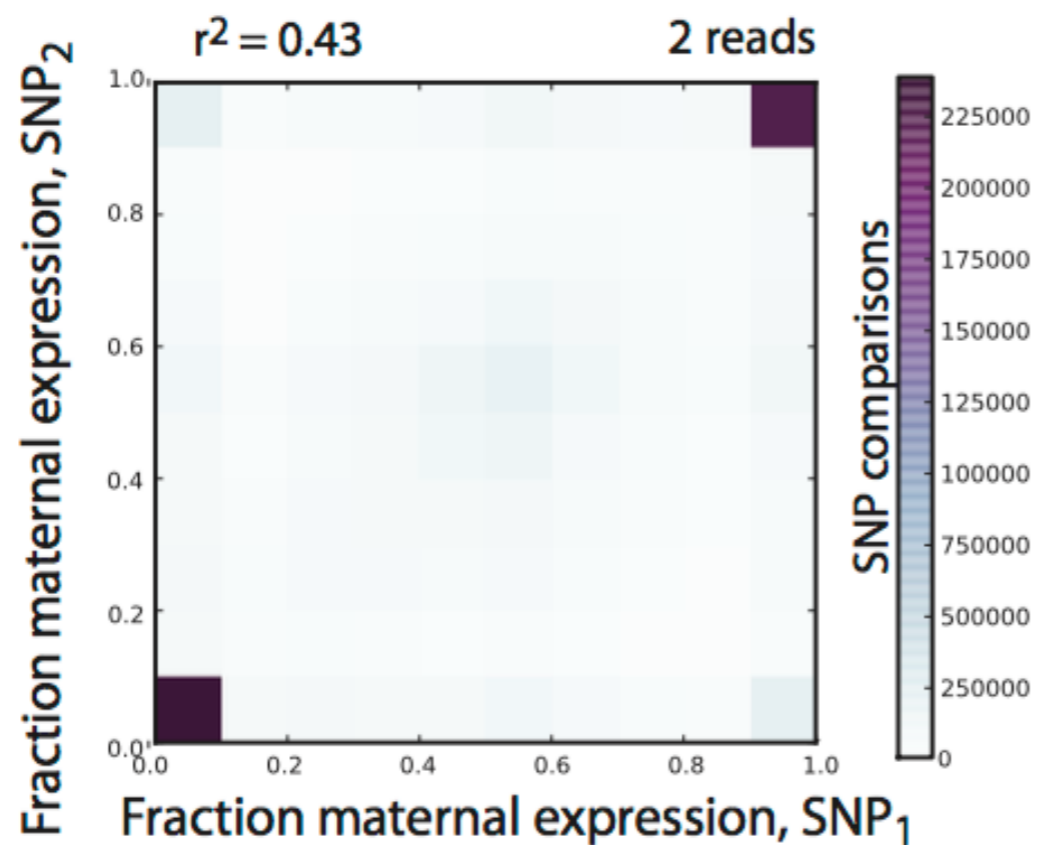
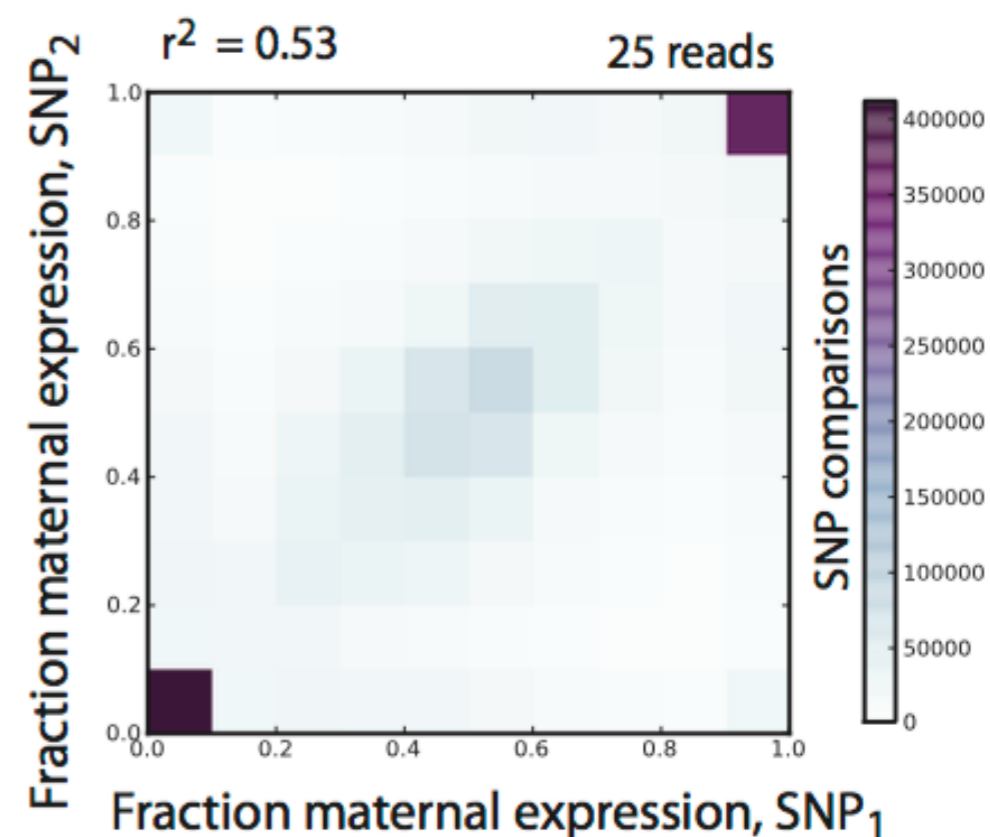
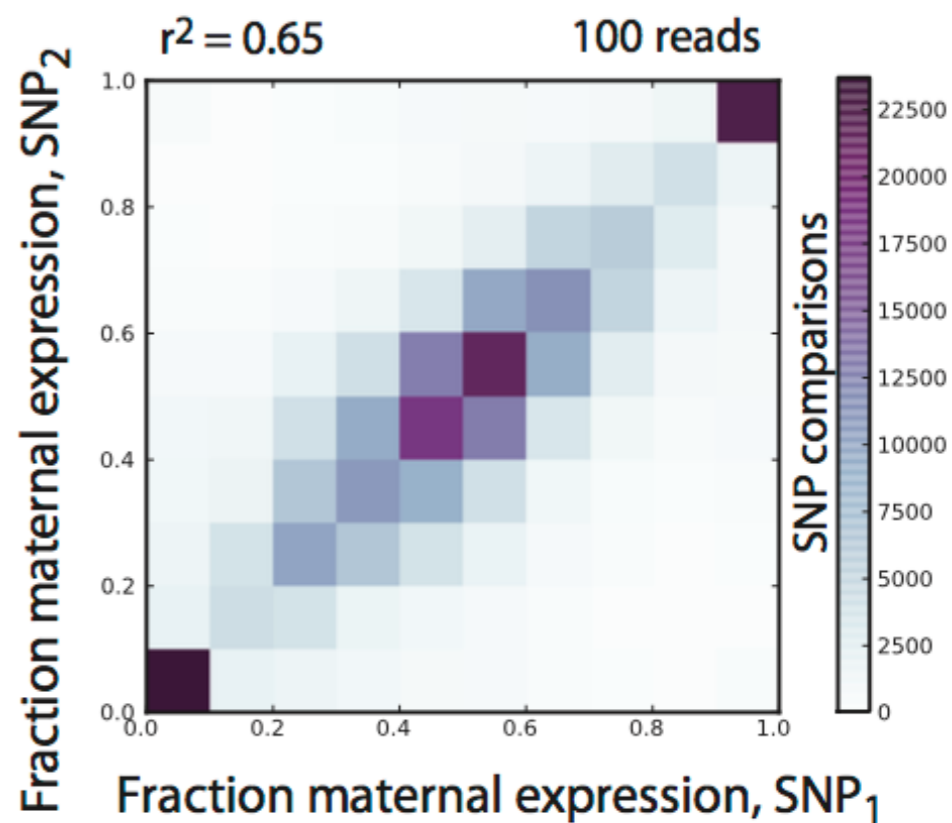




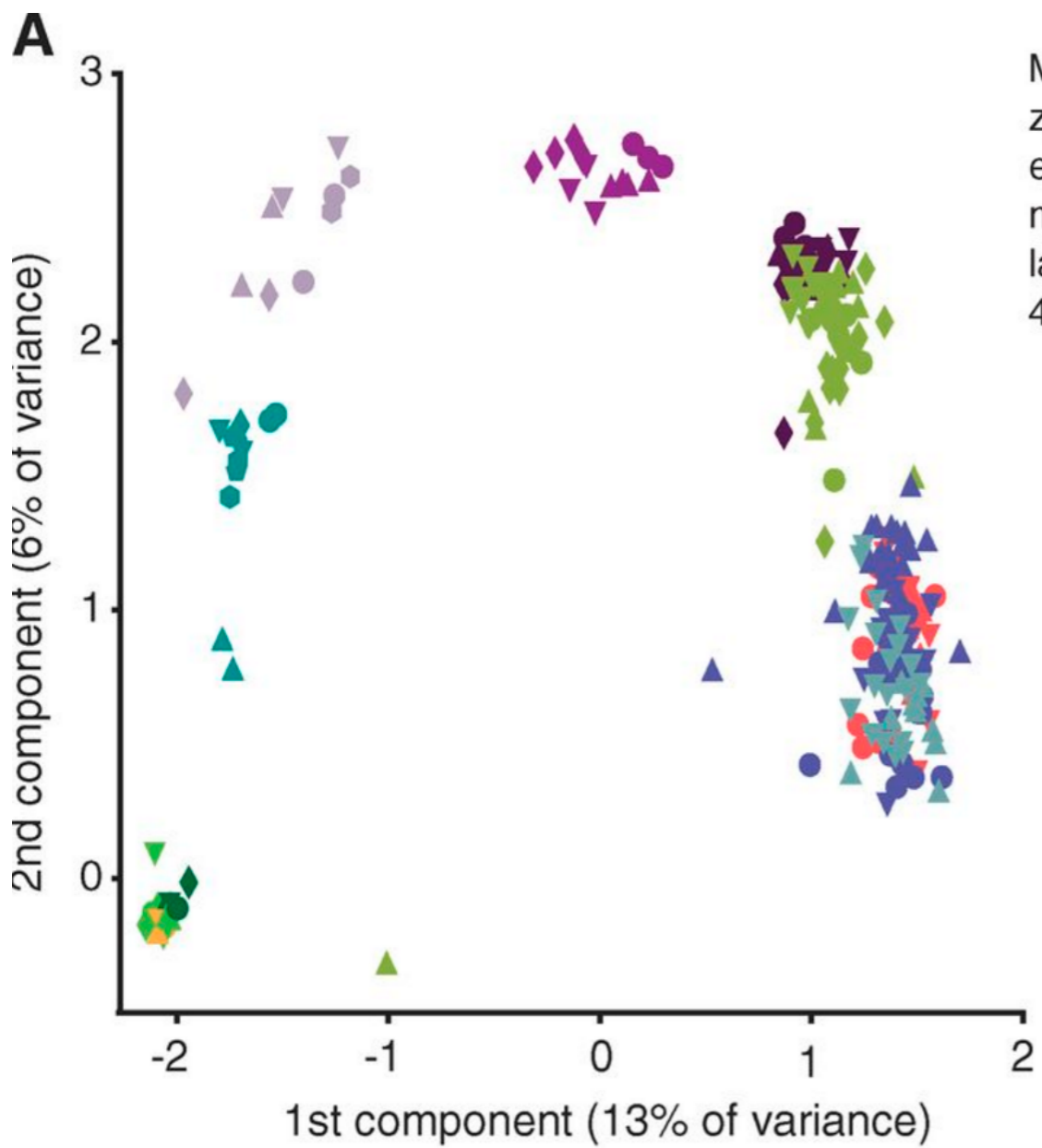
sc transcriptomes were highly reproducible with higher variability at later stages

82 % of all expressed ( $\geq 3$  RPKM in any cell) genes had at least one informative SNP

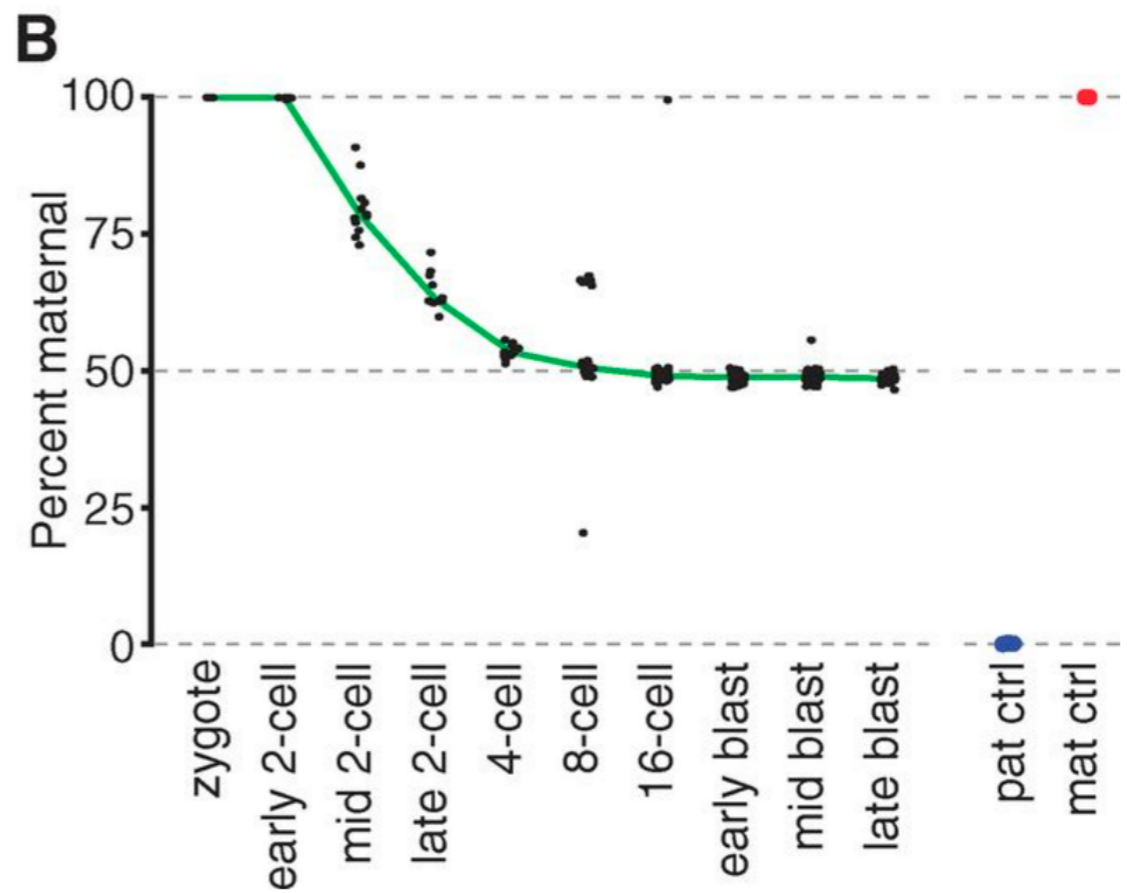


**A****B****C**

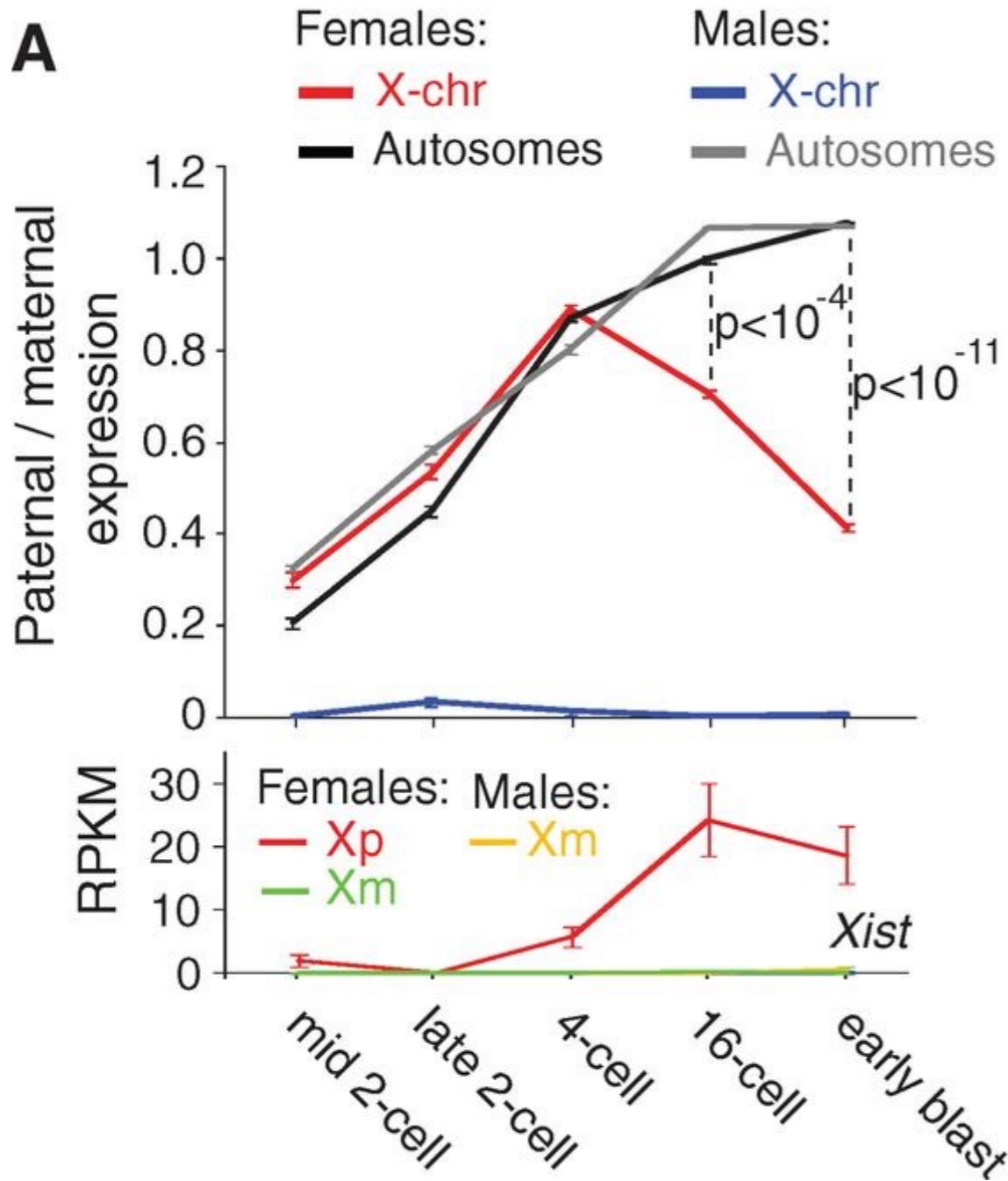
Fractions of maternal expression estimated from pairs of SNPs within the same genes ( $\geq 60$  bp apart)



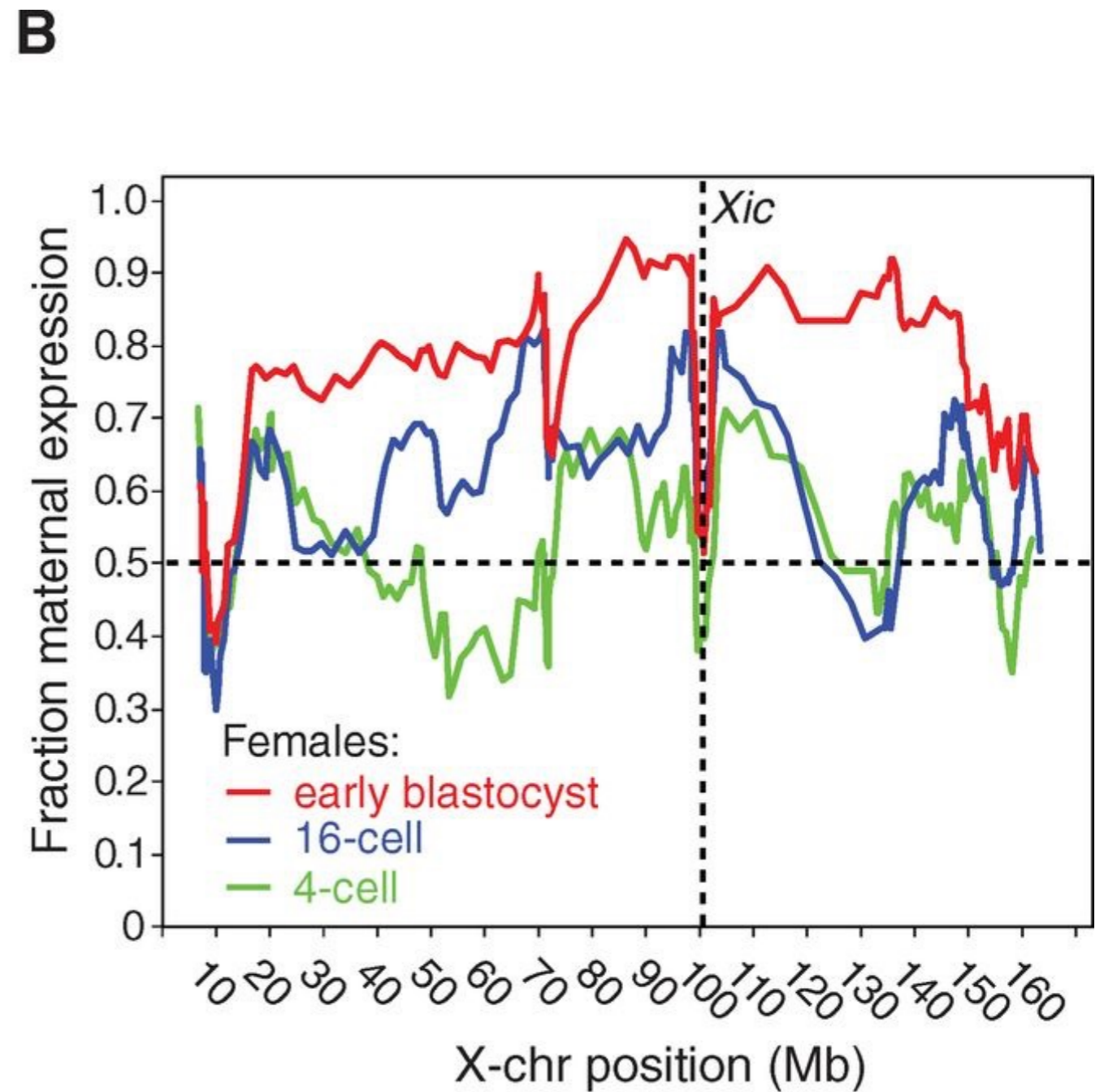
PCA clusters the cells stages effectively reconstructing the dynamics of PD



the mouse embryo inherits a large supply of maternal mRNA but that the bulk of this RNA is eliminated in the 2-cell embryo

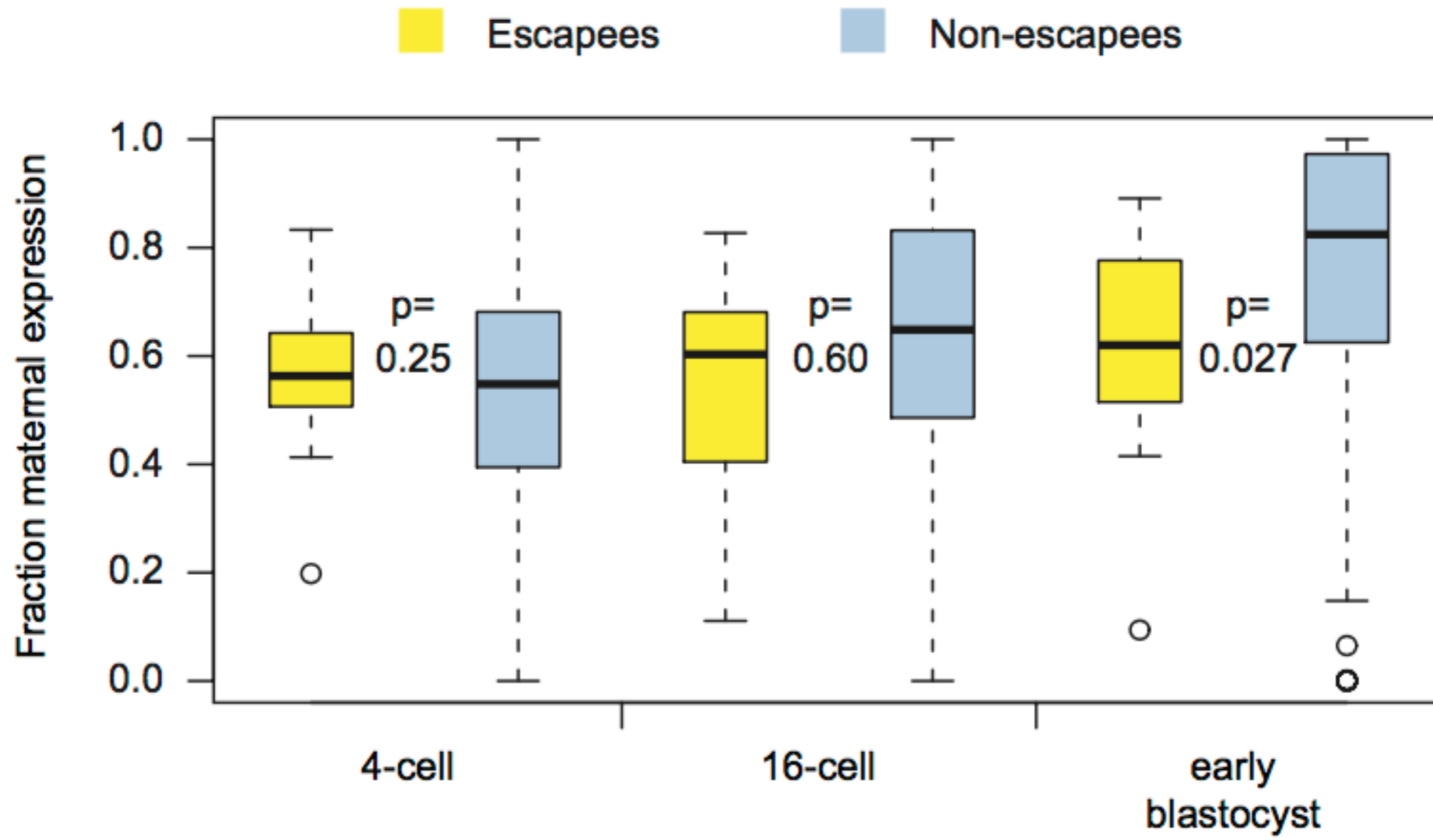


substantiating Xp inactivation rather than inheritance of pre-inactivated Xp



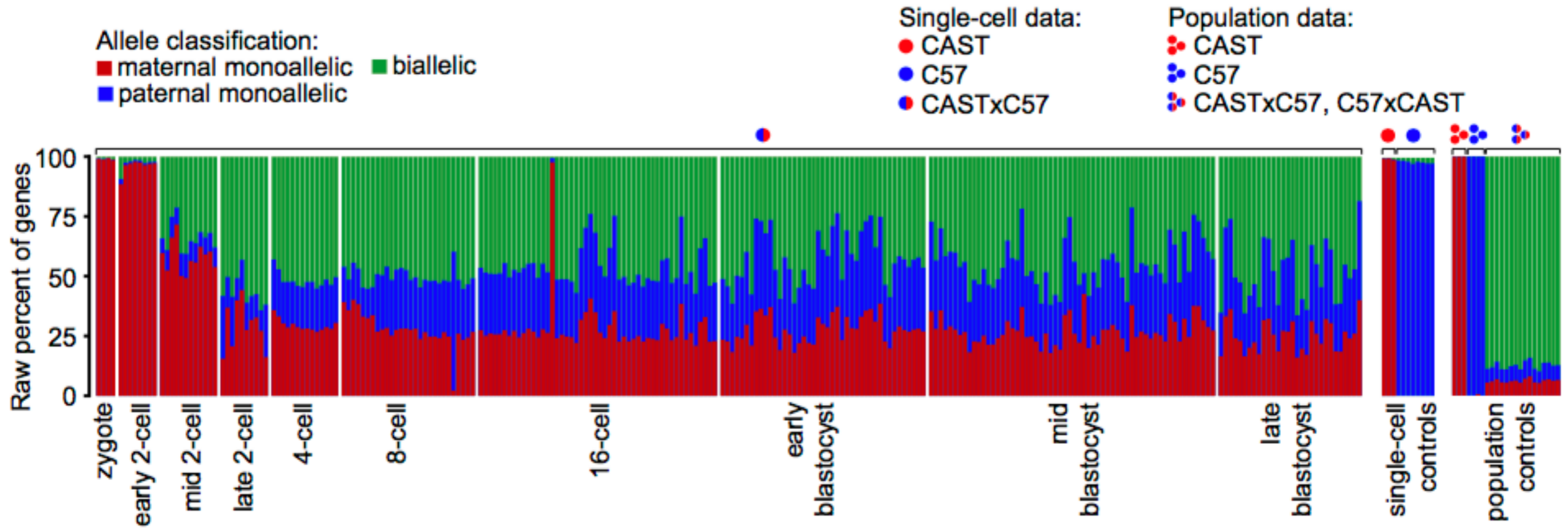
and that the spread of Xp silencing is not a simple function of the distance to Xic

'escapee' genes known from literature vs other X-chr genes

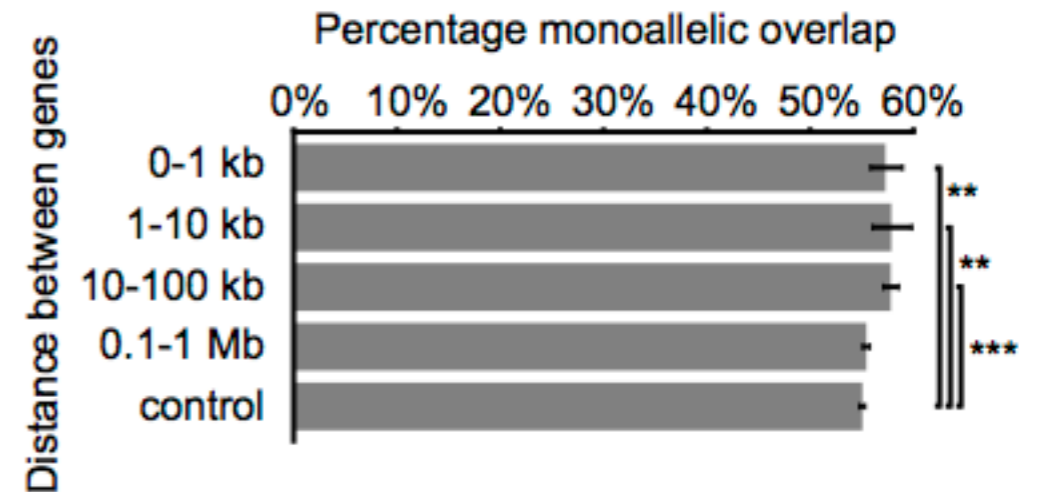




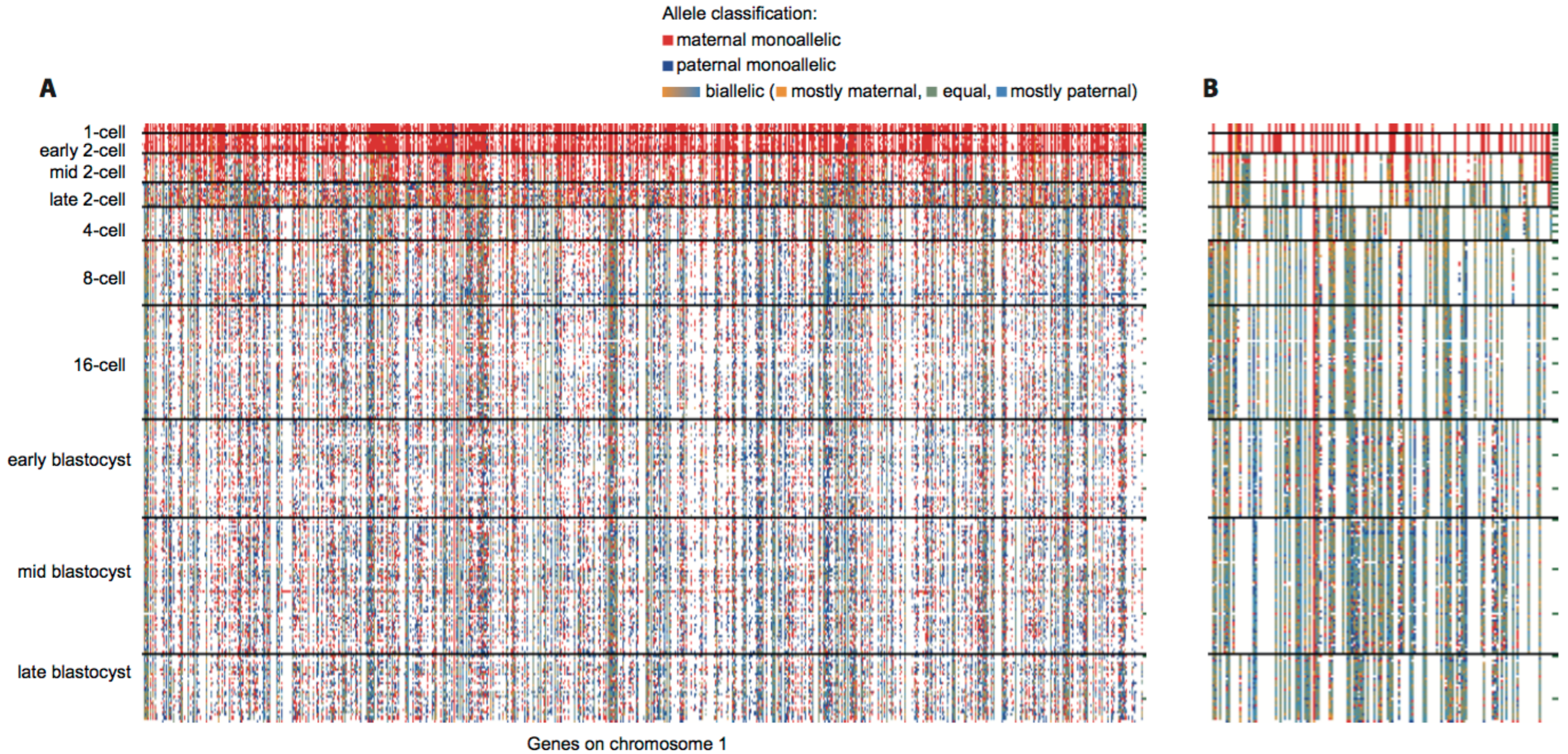
Analysis of raw monoallelic expressions in all cells and stages



Similarity between nearby monoallelically expressed genes, in how often they express the identical allele

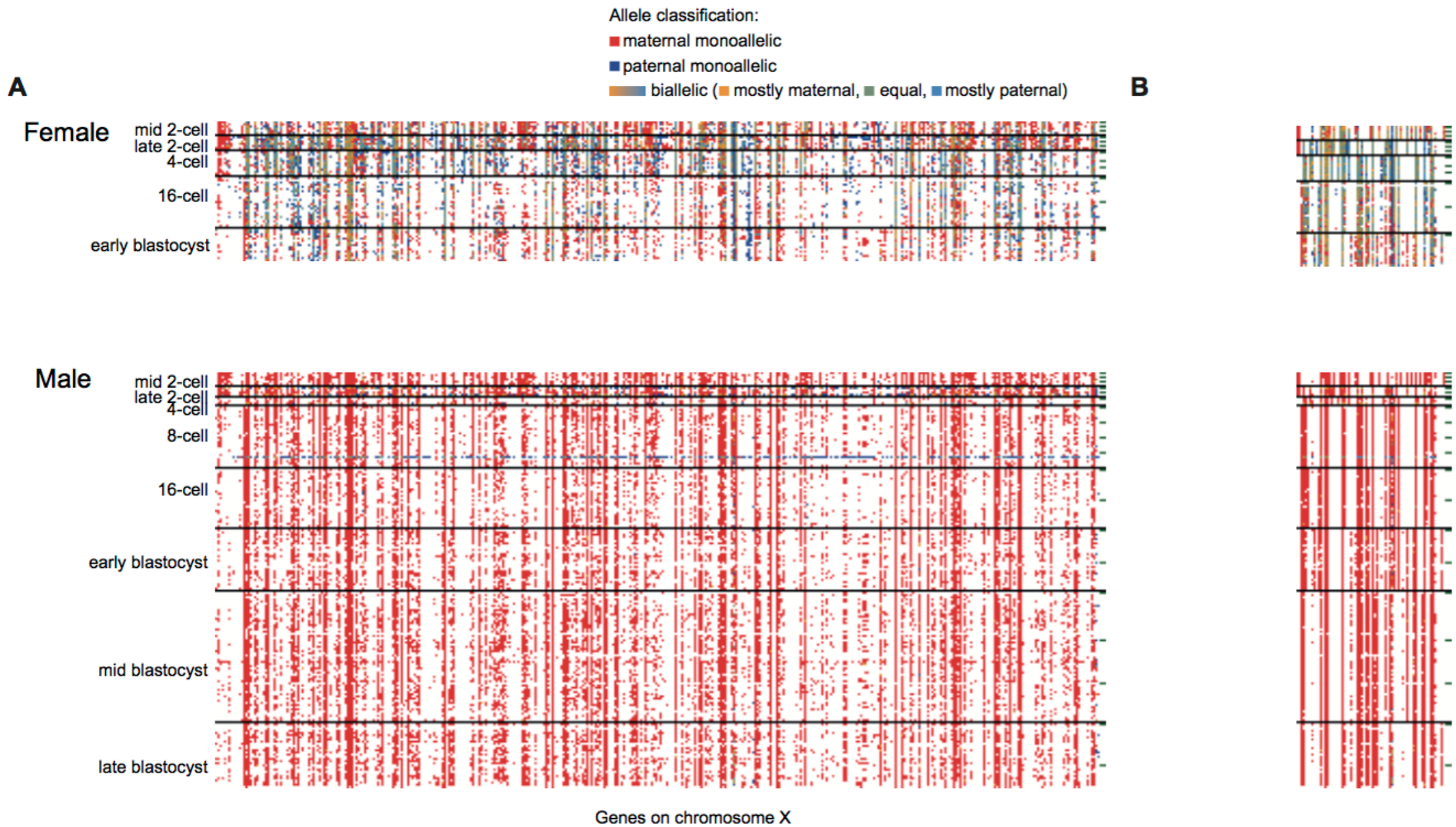


# Allelic transcription patterns



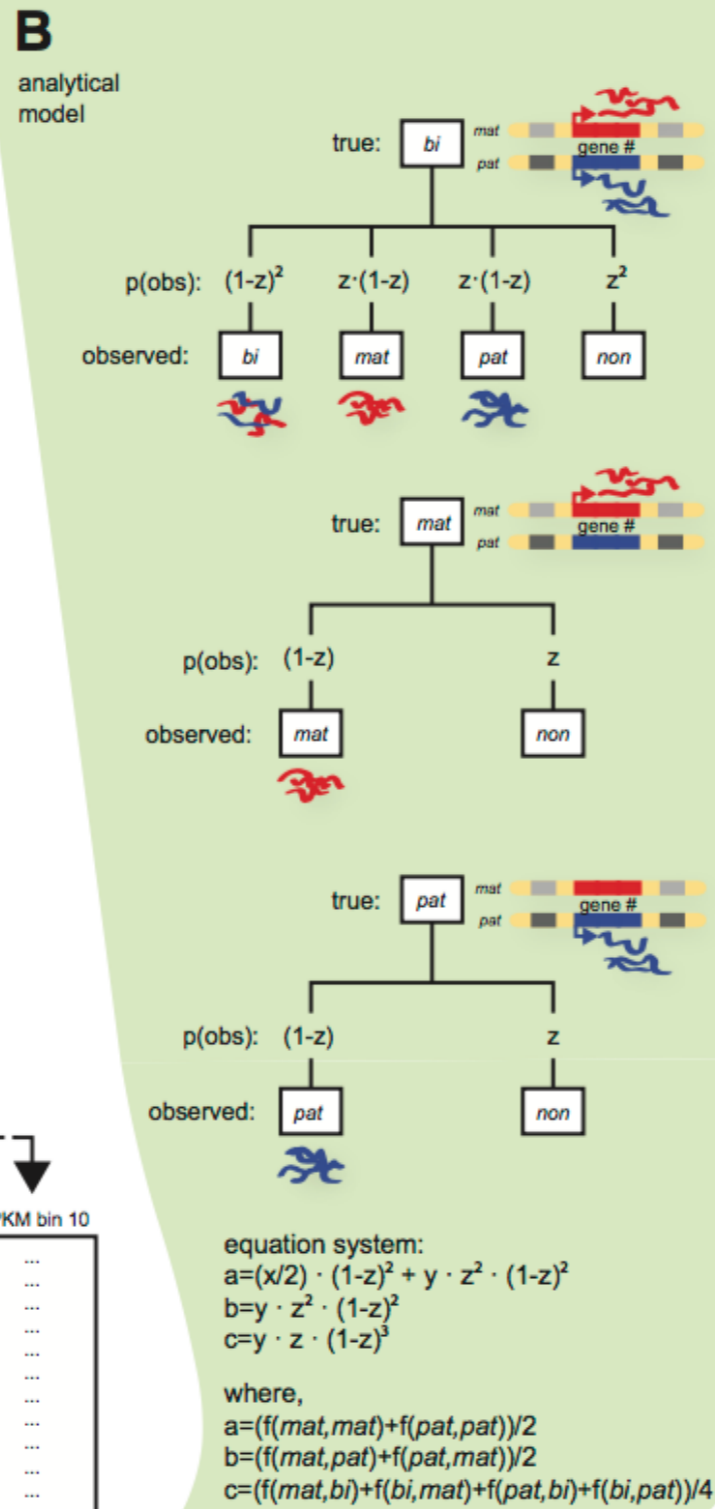
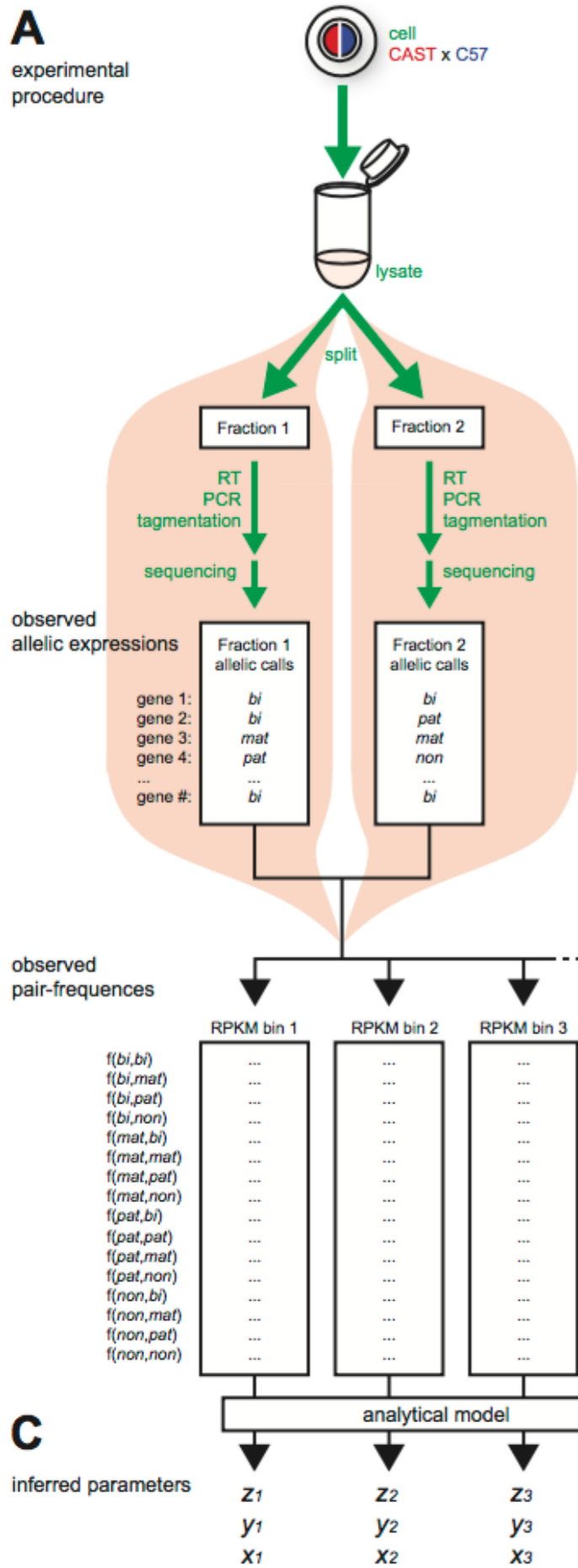
**Fig S5. Allelic transcription patterns for genes on chromosome 1**

Allele-specific gene expression for all cells (rows) and genes on chromosome 1 (columns), with genes sorted according to their genomic location. The allele expressions were classified into maternal monoallelic (red), paternal monoallelic (blue) and different degrees of biallelic expression (see color legend). (A) shows the raw monoallelic expression patterns and (B) is filtered for genes expressed above 90 RPKM.

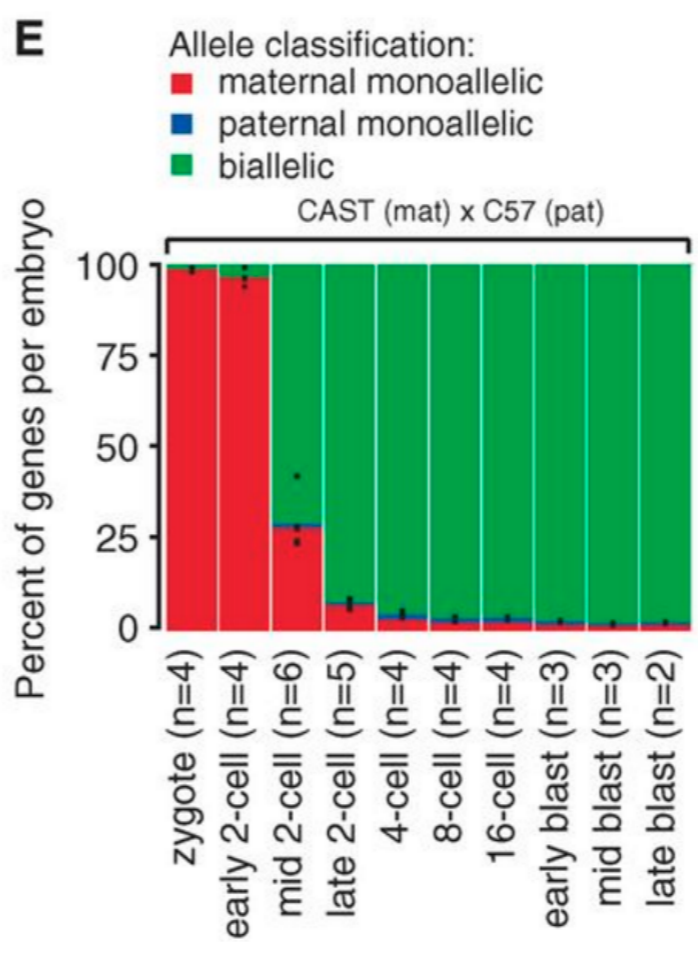
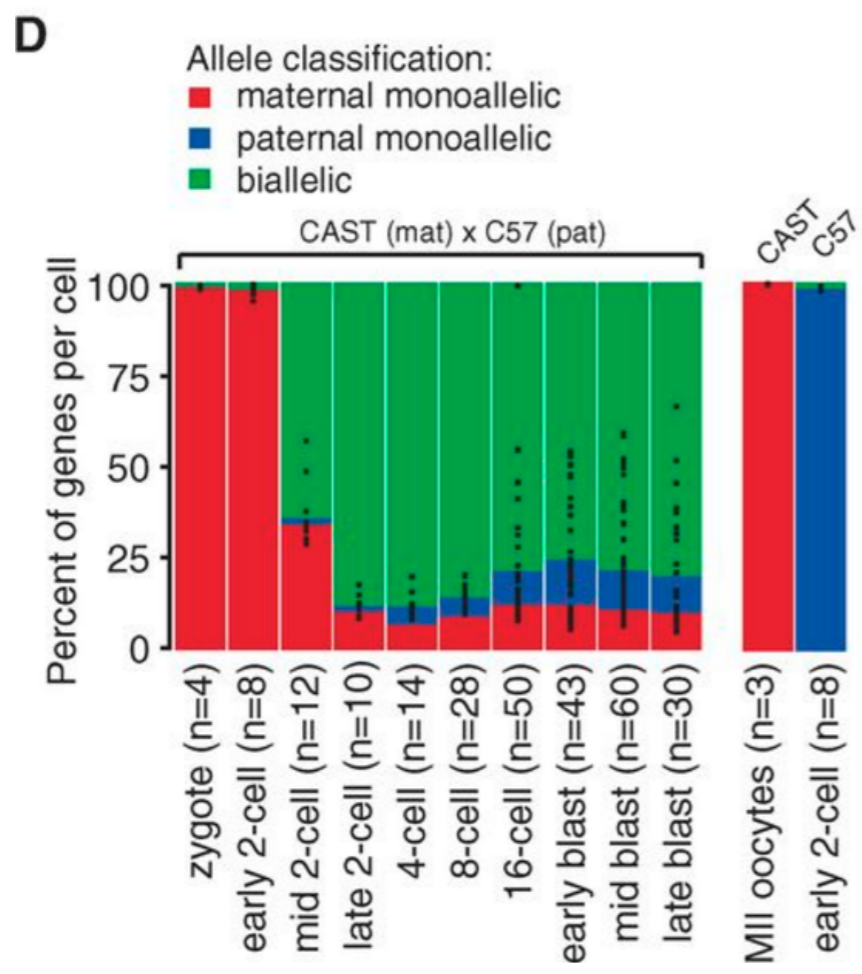
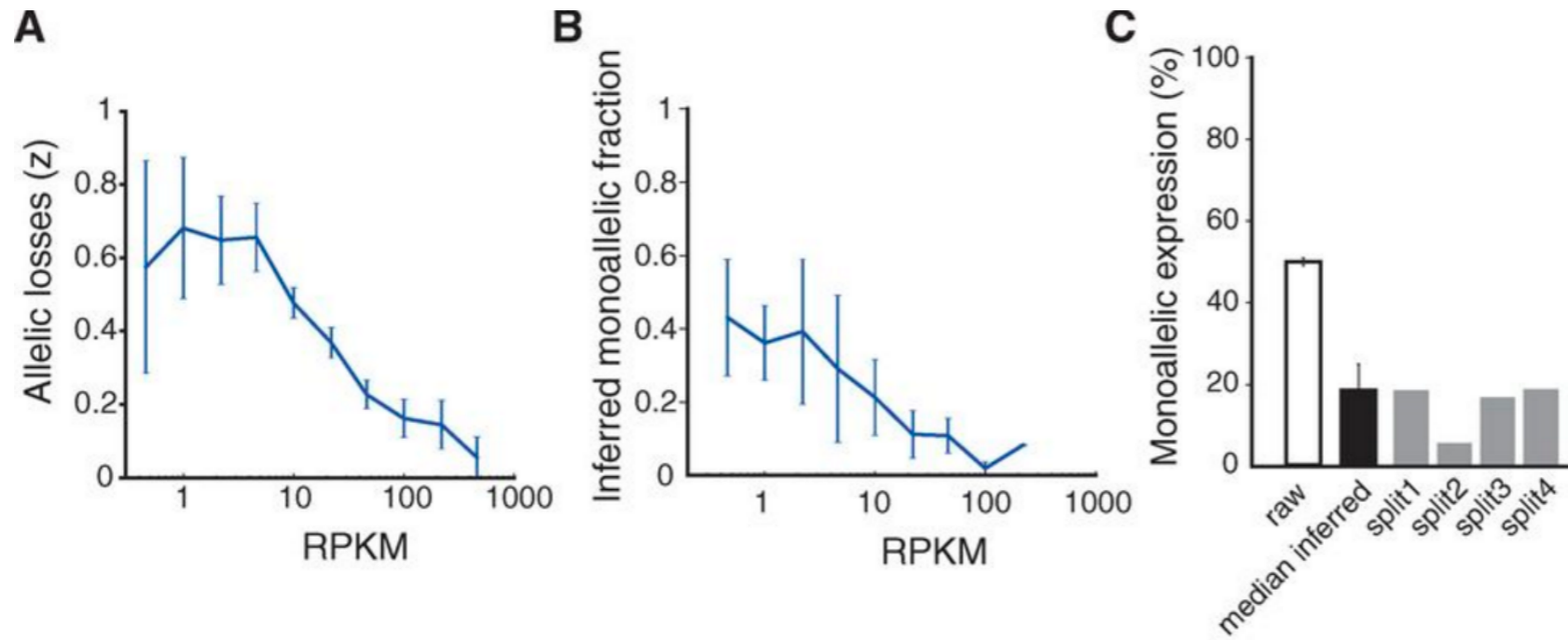


**Fig S24. Allelic transcription patterns for genes on chromosome X**

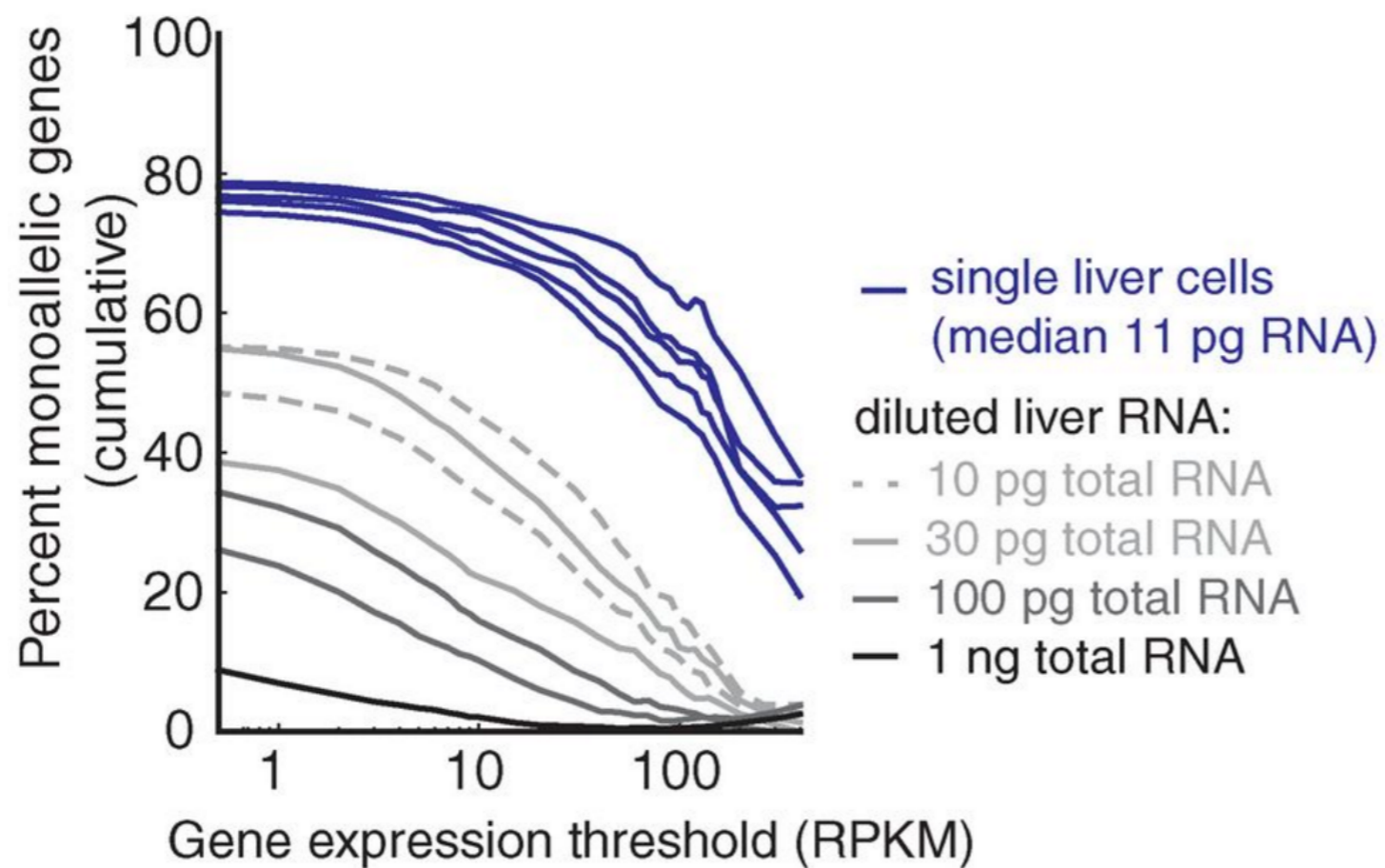
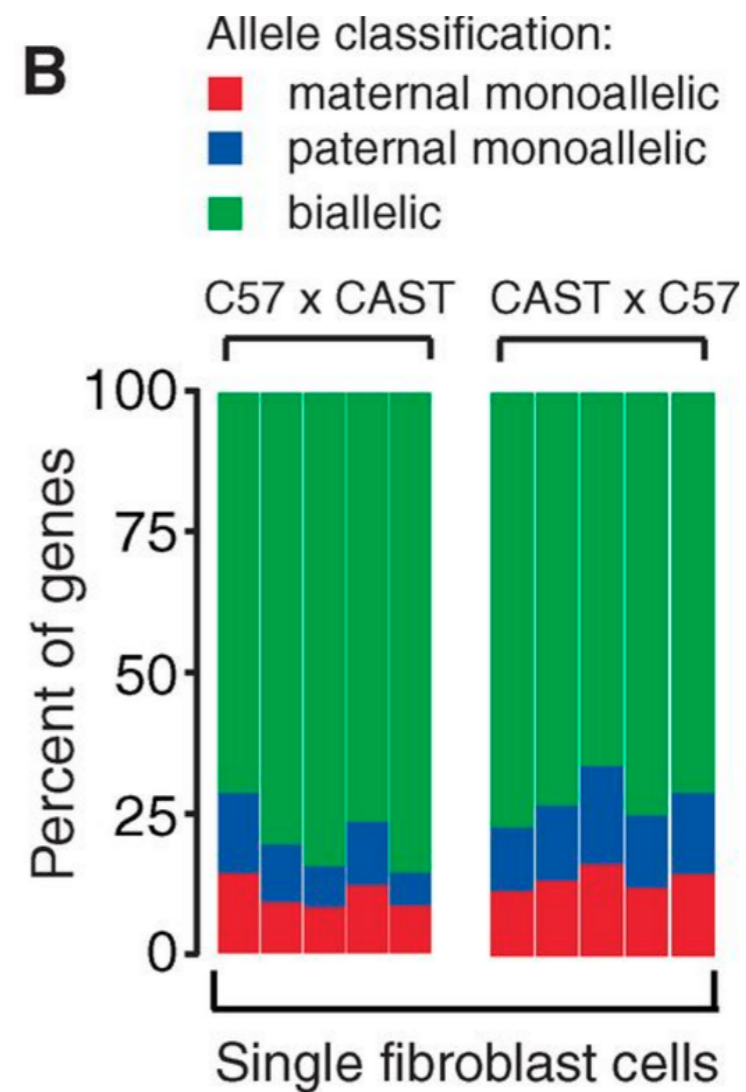
Allele-specific gene expression for all cells (rows) and genes on chromosome X (columns), with genes sorted according to their genomic location. The allele expressions were classified into maternal monoallelic (red), paternal monoallelic (blue) and different degrees of biallelic expression (see color legend). (A) shows the raw monoallelic expression patterns and (B) is filtered (per stage) for genes expressed above 90 RPKM.



Estimation of stochastic losses in sc RNA experiment (due to limitations in detection and quantification of two alleles for weakly transcribed genes)



Pooling cells by embryo removes essentially all mono allelic expression - high degree of cell-specific randomness

**A****B**

Article

## Biased Allelic Expression in Human Primary Fibroblast Single Cells

Christelle Borel<sup>1, 8</sup>, Pedro G. Ferreira<sup>1, 2, 3, 8</sup>, Federico Santoni<sup>1</sup>, Olivier Delaneau<sup>1, 2, 3</sup>, Alexandre Fort<sup>4</sup>, Konstantin Y. Popadin<sup>1</sup>, Marco Garieri<sup>1</sup>, Emilie Falconnet<sup>1</sup>, Pascale Ribaux<sup>1</sup>, Michel Guipponi<sup>1, 5</sup>, Ismael Padioleau<sup>1</sup>, Piero Carninci<sup>4</sup>, Emmanouil T. Dermitzakis<sup>1, 2, 3, 6, 7</sup>,  , Stylianos E. Antonarakis<sup>1, 2, 5</sup>,  

<sup>1</sup> Department of Genetic Medicine and Development, University of Geneva, 1211 Geneva, Switzerland

<sup>2</sup> Institute of Genetics and Genomics of Geneva, 1211 Geneva, Switzerland

<sup>3</sup> Swiss Institute of Bioinformatics, 1211 Geneva, Switzerland

<sup>4</sup> Division of Genomic Technologies, RIKEN Center for Life Science Technologies, Yokohama, Kanagawa 230-0045, Japan

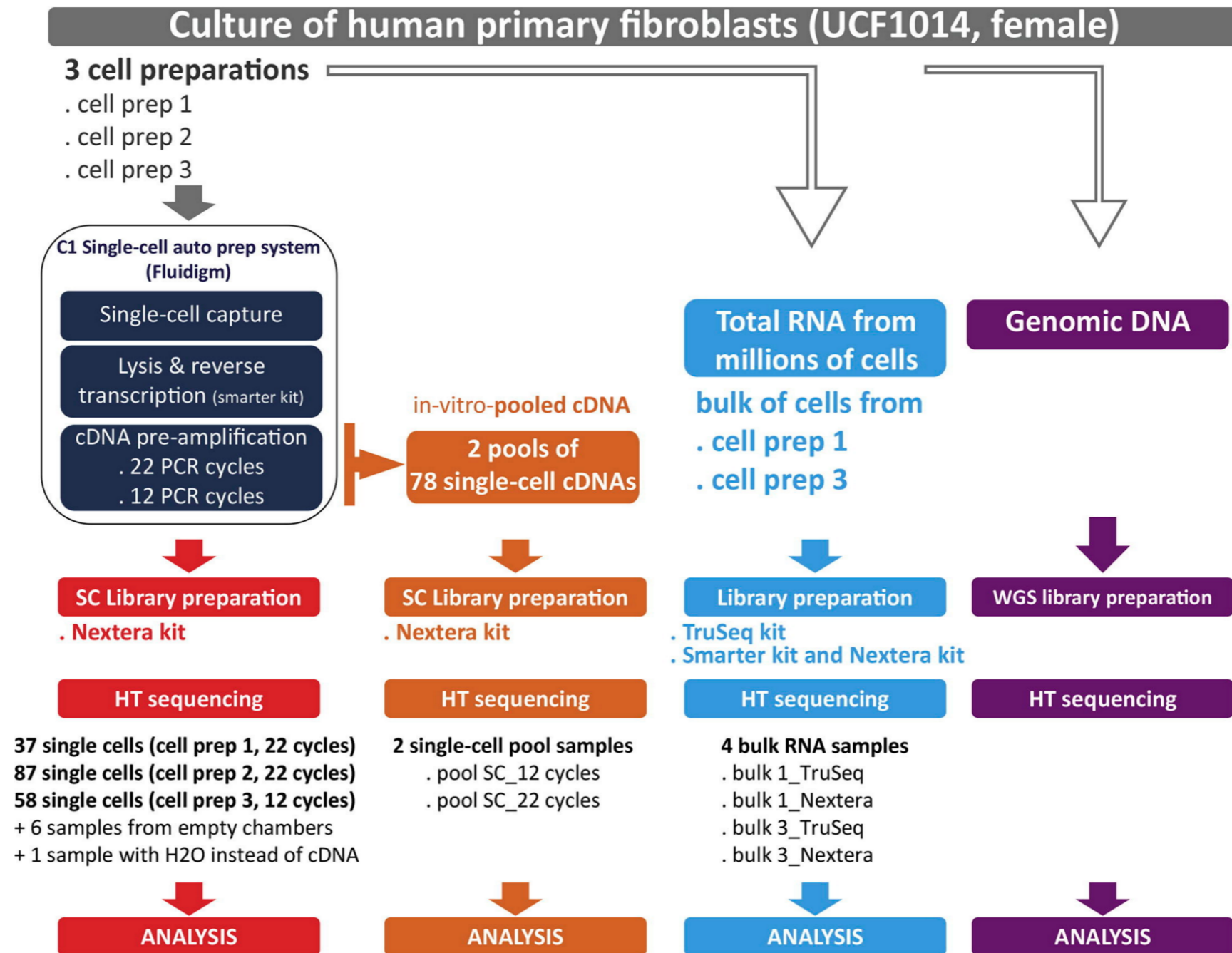
<sup>5</sup> Service of Genetic Medicine, University Hospitals of Geneva, 1211 Geneva, Switzerland

<sup>6</sup> Center of Excellence for Genomic Medicine Research, King Abdulaziz University, Jeddah 21589, Saudi Arabia

<sup>7</sup> Biomedical Research Foundation Academy of Athens, Athens 11527, Greece

- scRNA-seq and bcRNA-seq

- 203 single cells (163 female, UCF1014, cells and 40 human fetal primary fibroblast culture, T2N, cells) from two different human primary fibroblast cell lines





# Culture of human primary fibroblasts (T2N, female)

## 1 cell preparation

C1 Single-cell auto prep system  
(Fluidigm)

Single-cell capture

Lysis & reverse  
transcription (smarter kit)

cDNA pre-amplification  
. 22 PCR cycles

SC Library preparation

. Nextera kit

HT sequencing

40 single cells (22 cycles)

ANALYSIS

Total RNA from  
millions of cells

Library preparation

. TruSeq kit

HT sequencing

1 bulk RNA sample  
. bulk 1\_Trueq

ANALYSIS

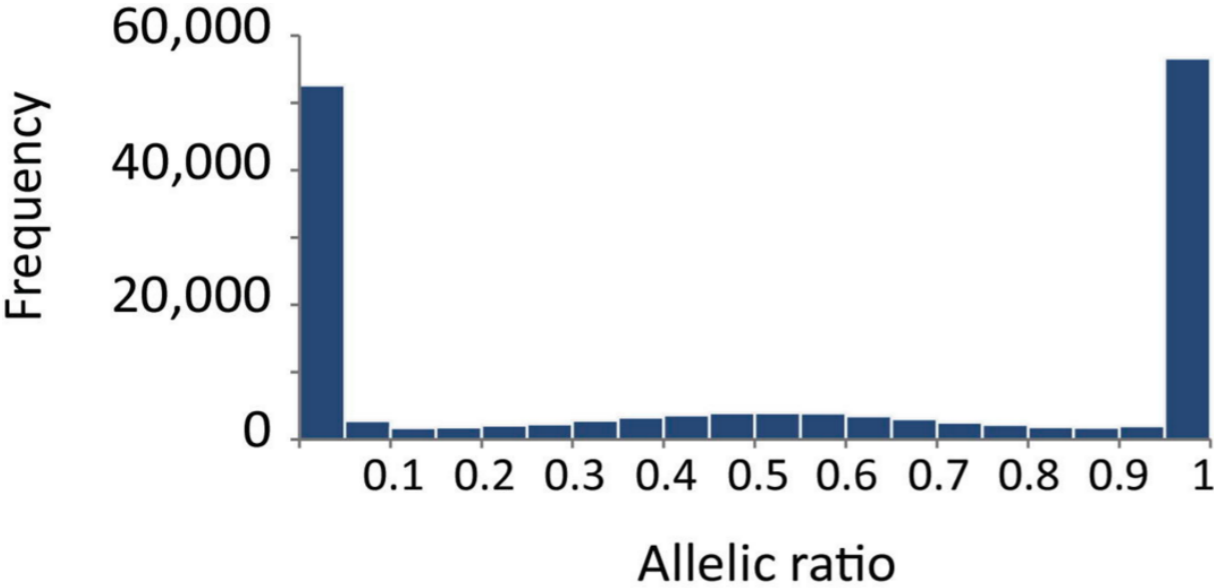
Genomic DNA

WGS library preparation

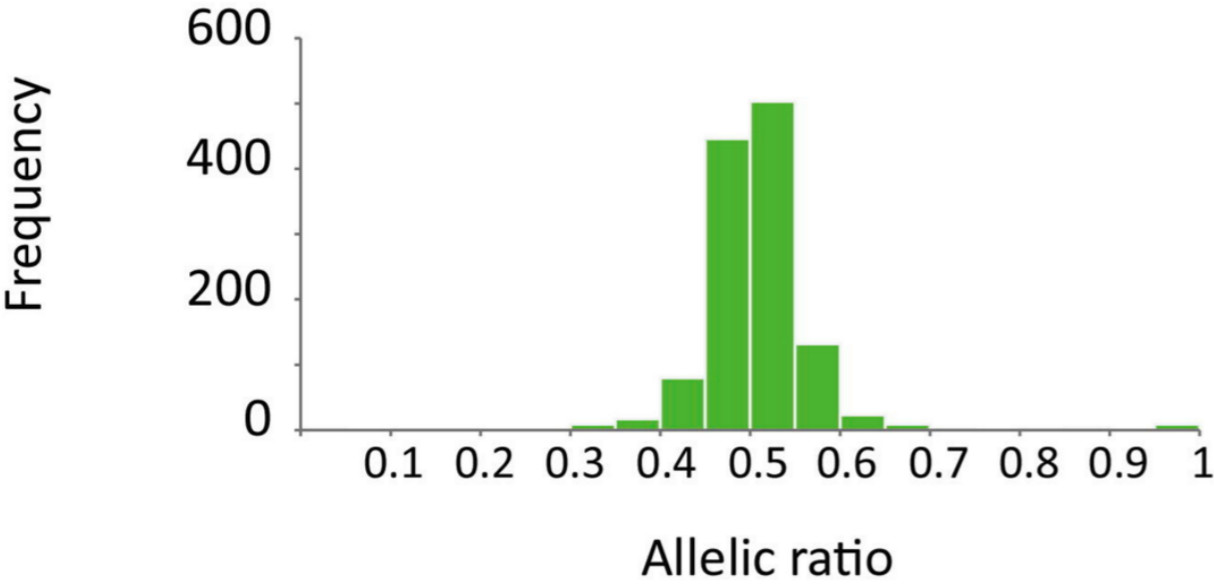
HT sequencing

ANALYSIS

Frequency distribution of the allelic ratios of 35,763 hetSNPVs ( $\geq 20$  RPSM) in 163 single cells, bulk cell samples, and the pool of single-cell samples



SINGLE CELLS (UCF1014, n = 163)

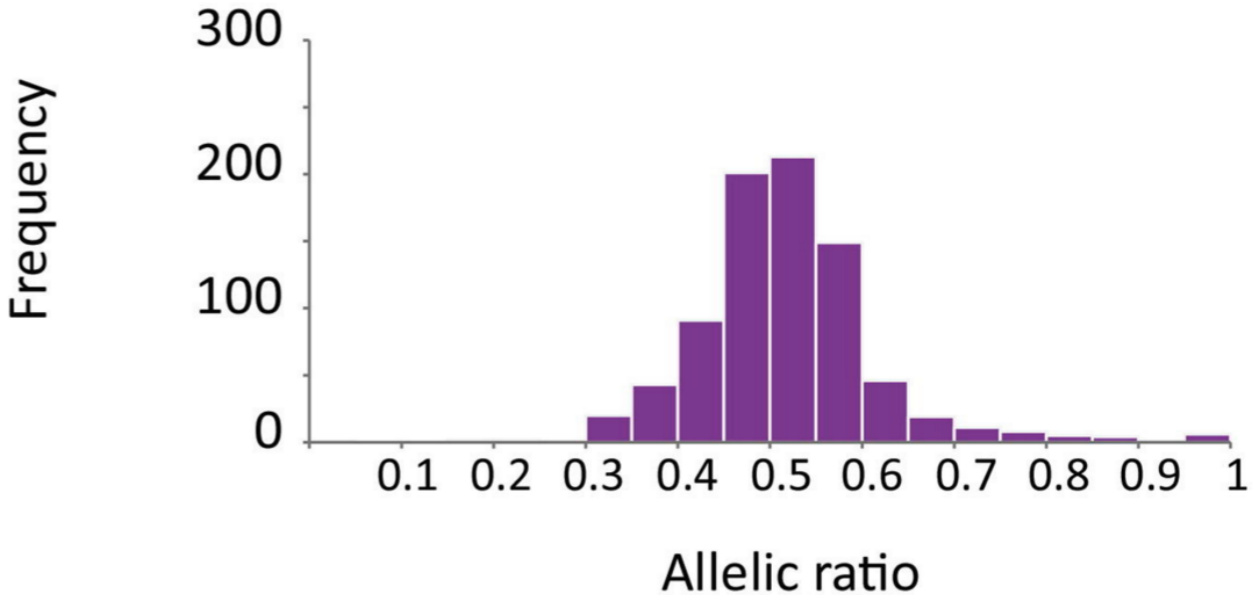


BULK OF CELLS (UCF1014, n = 4)

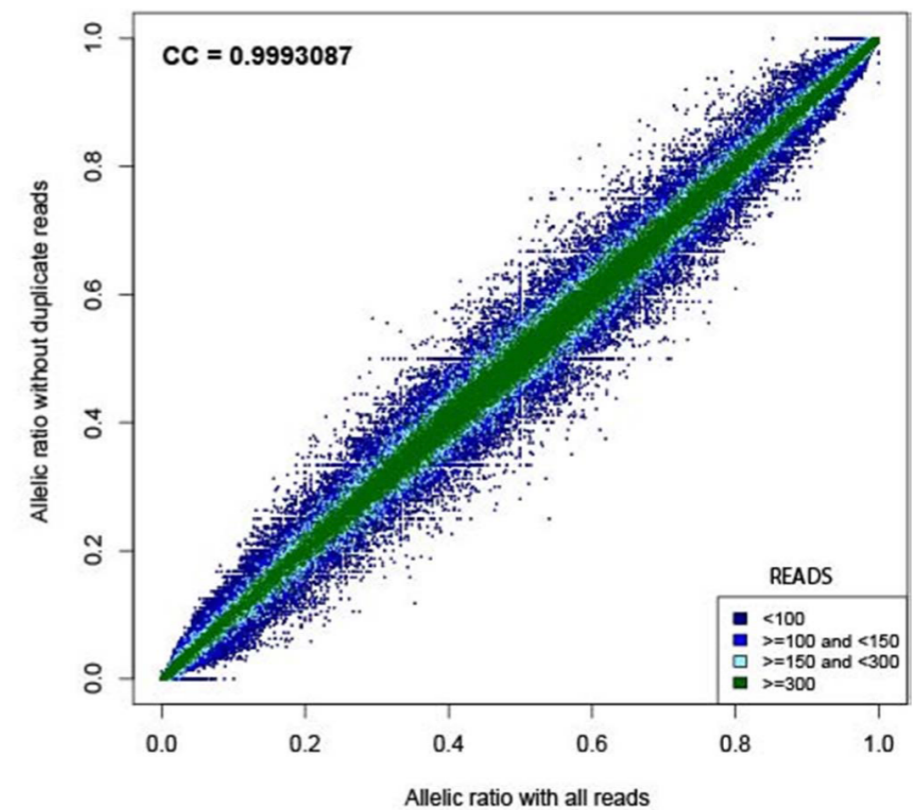
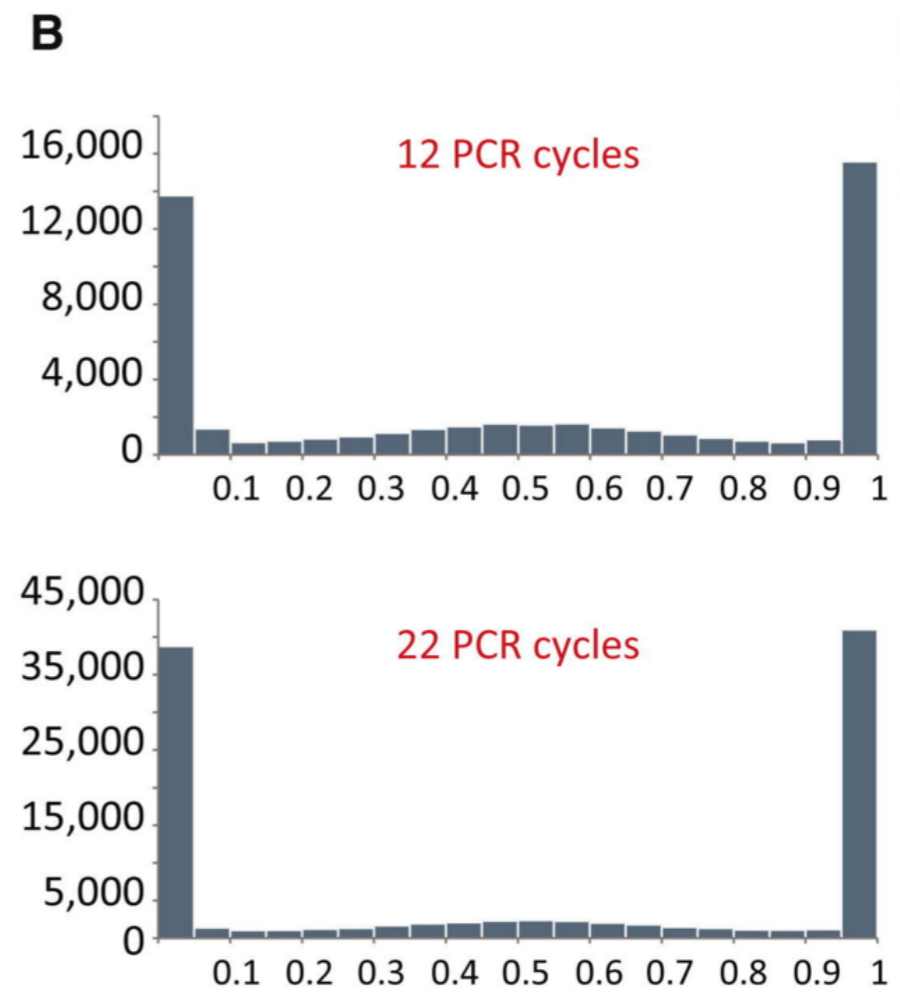
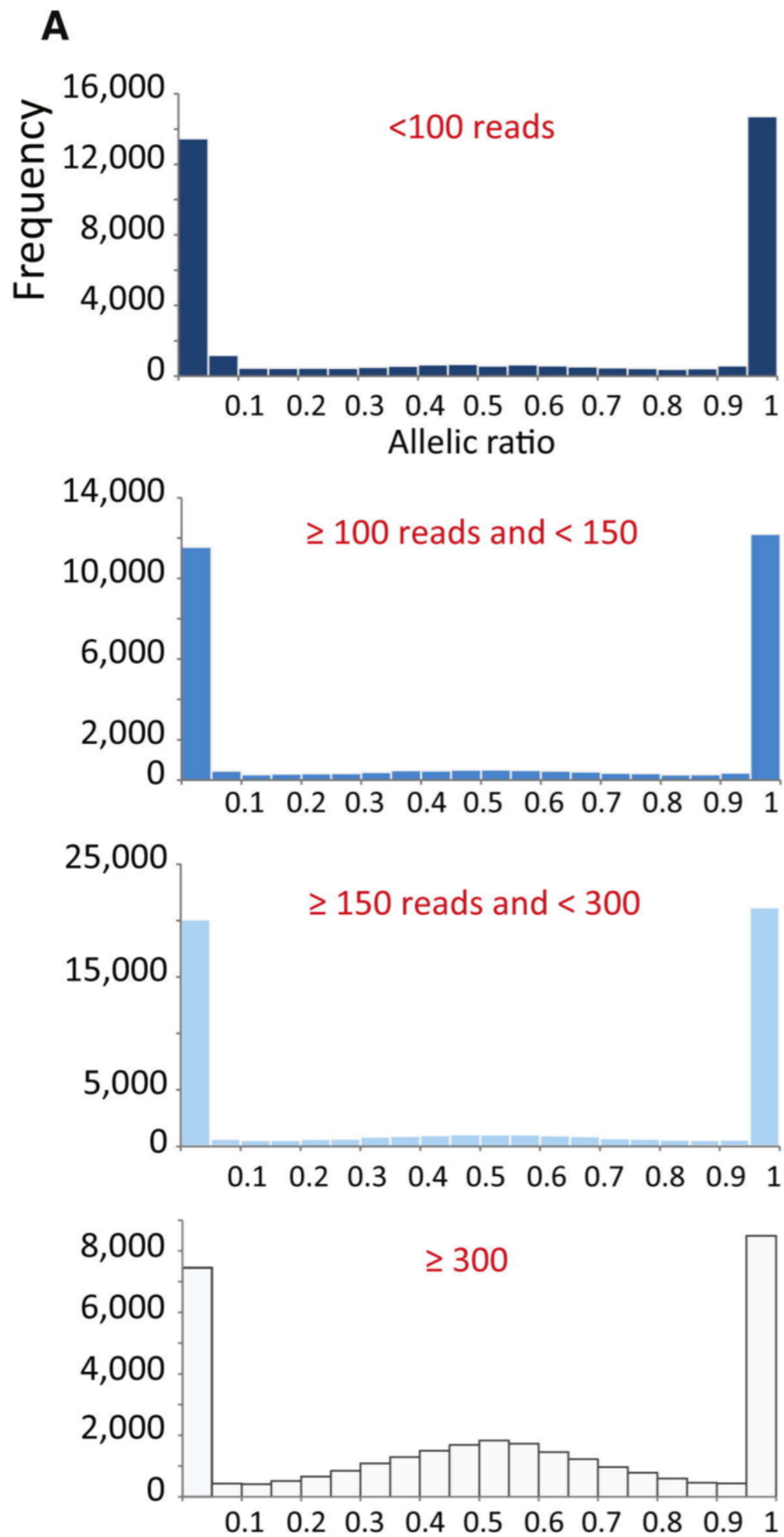
23.6 % of the hetSNVs had AR in 0.2 - 0.8 in sc

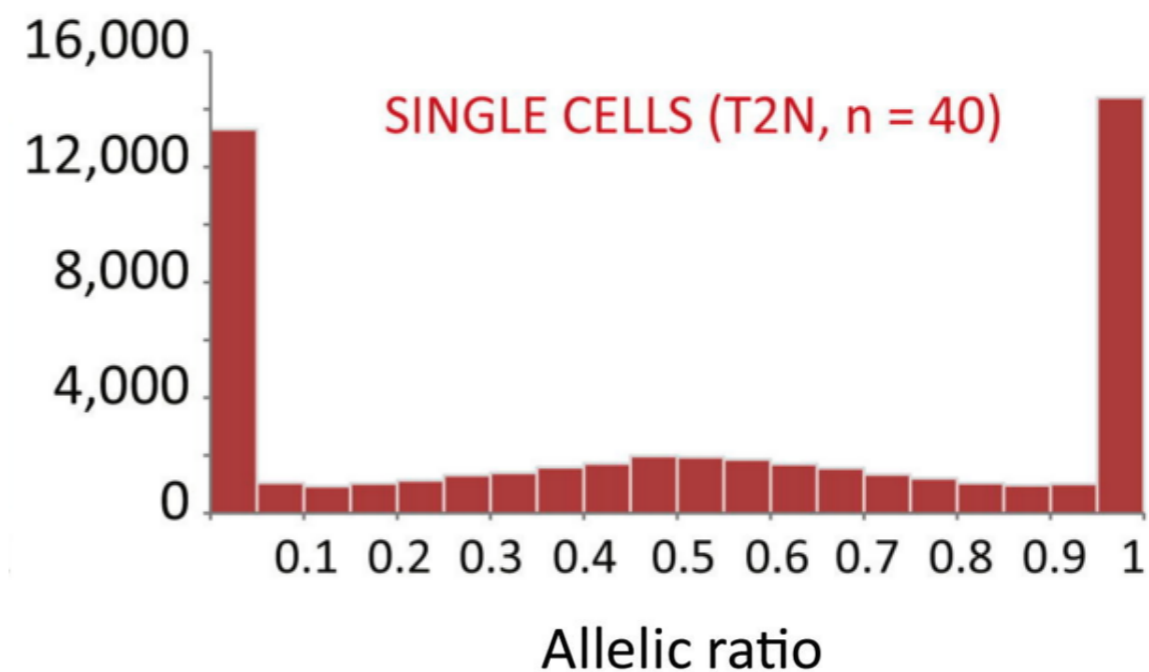
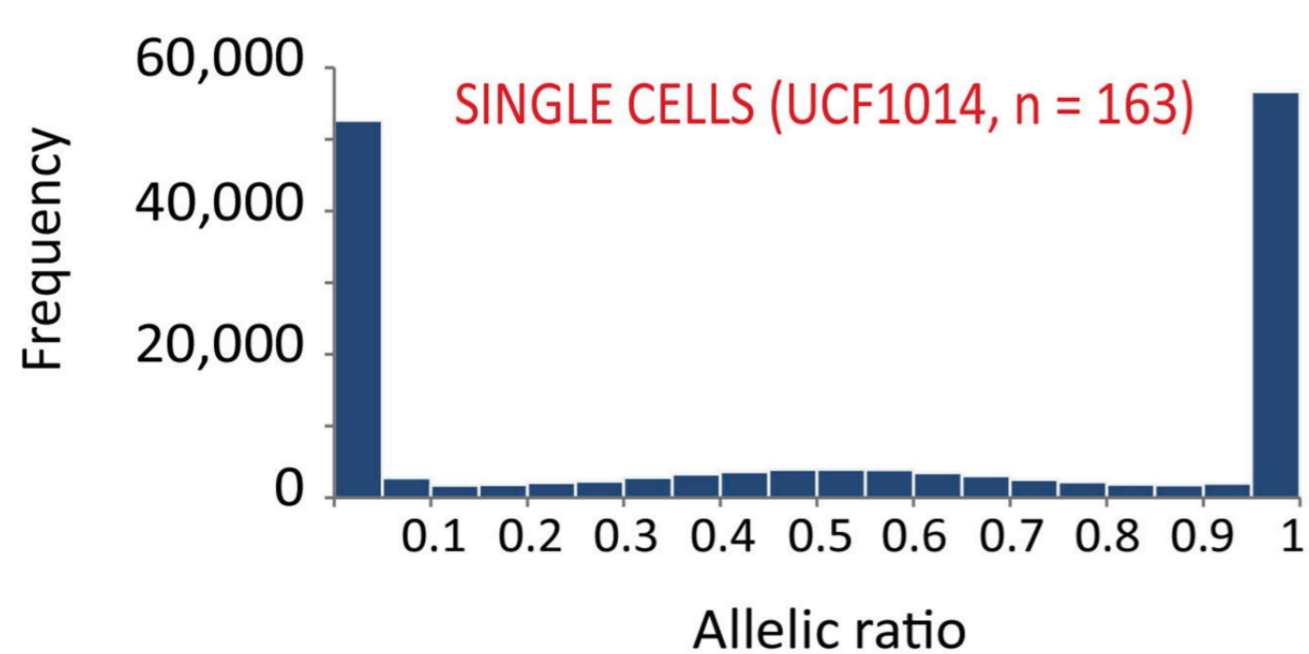
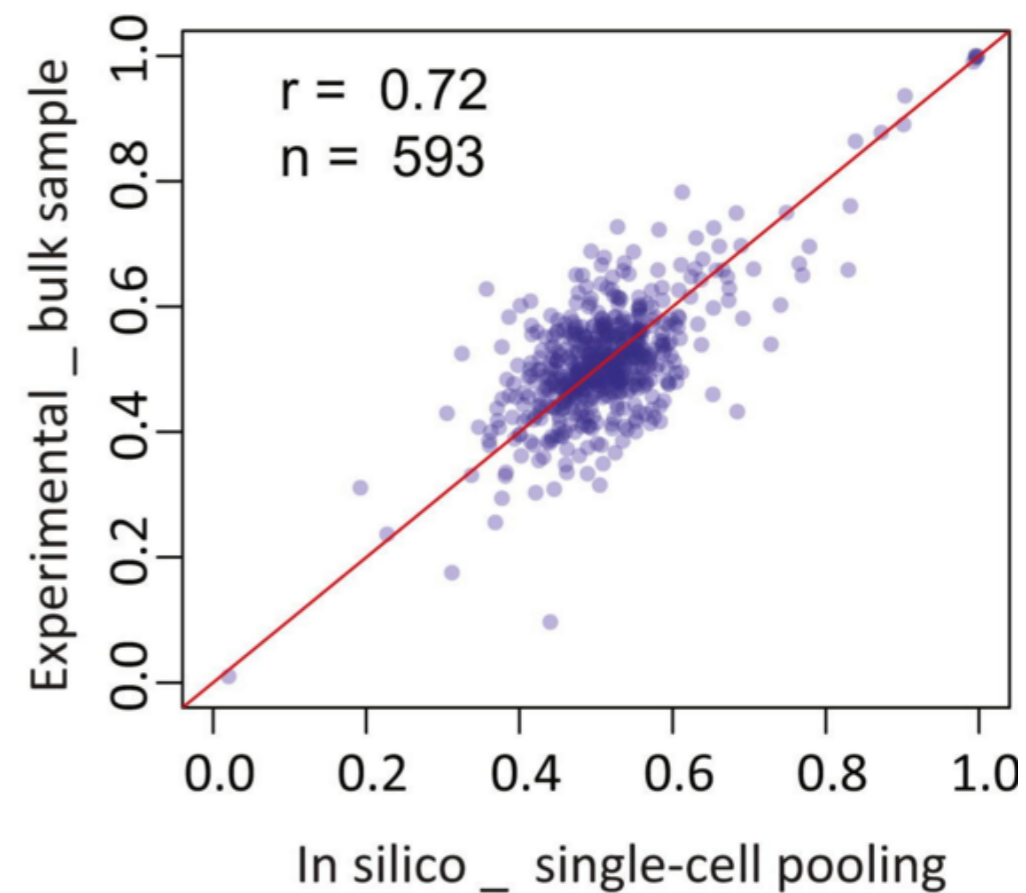
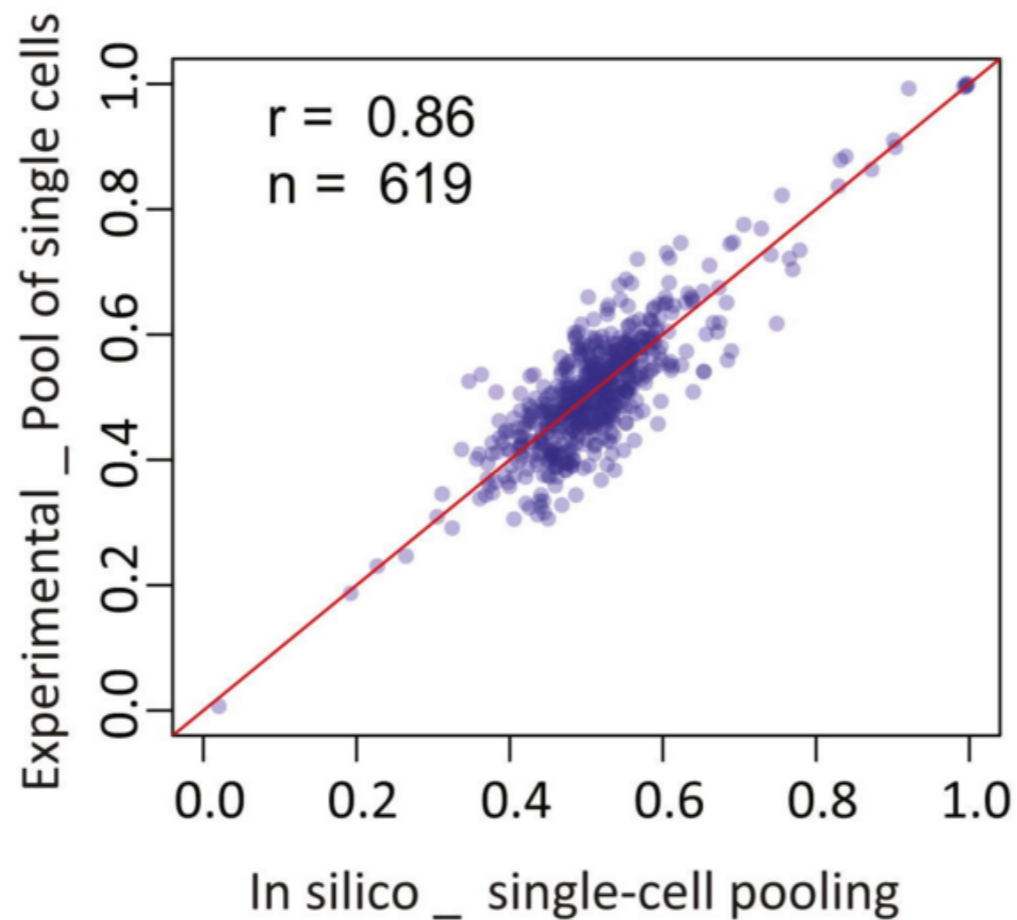
98.3 % of 545 hetSNVs w RPSM  $\geq 20$  had AR 0.2-0.8 in bc samples

97.7 % of 480 hetSNVs w RPSM  $\geq 20$  had AR 0.2-0.8 in sc pools



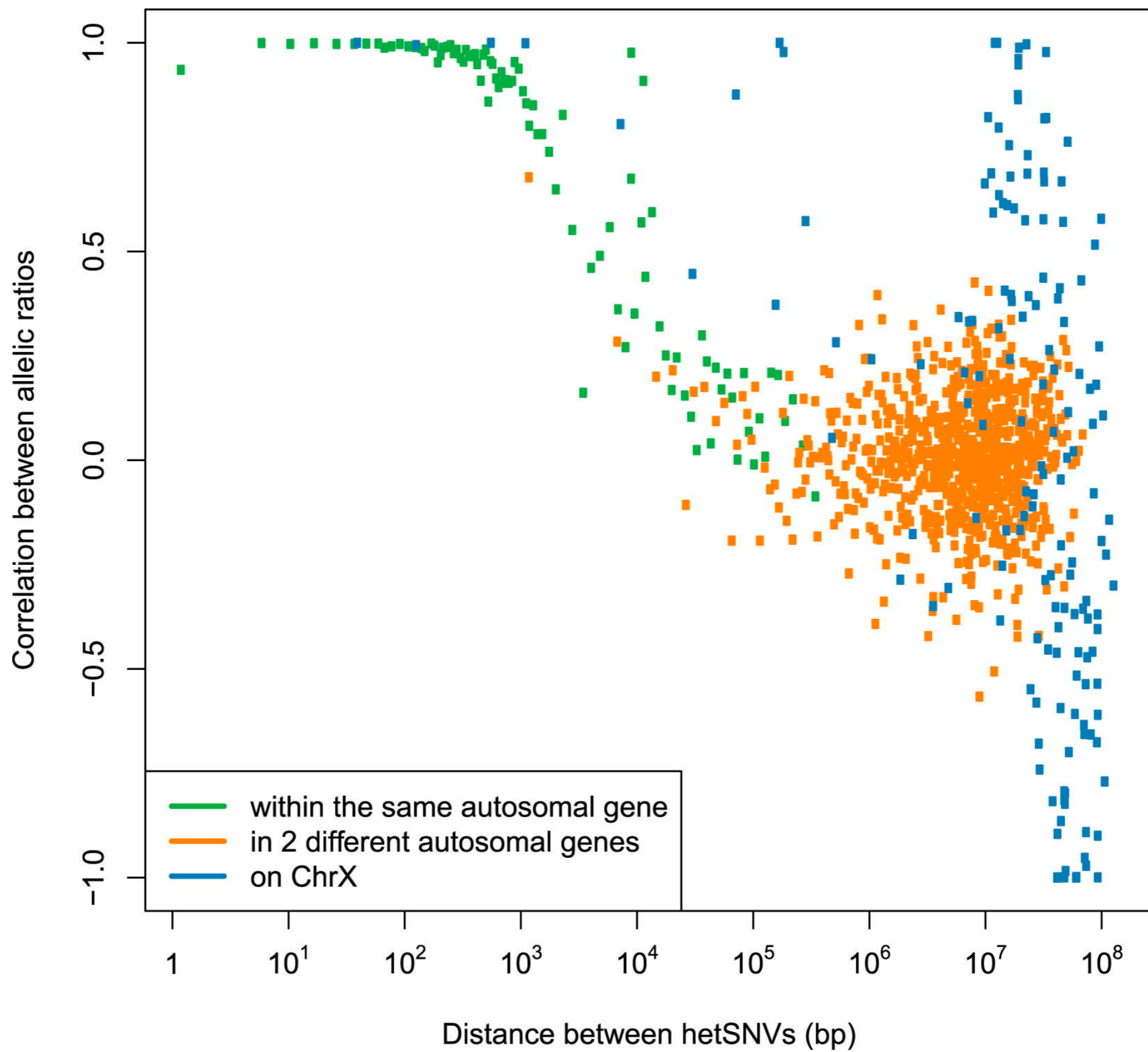
POOL OF CELLS (UCF1014, n = 2)

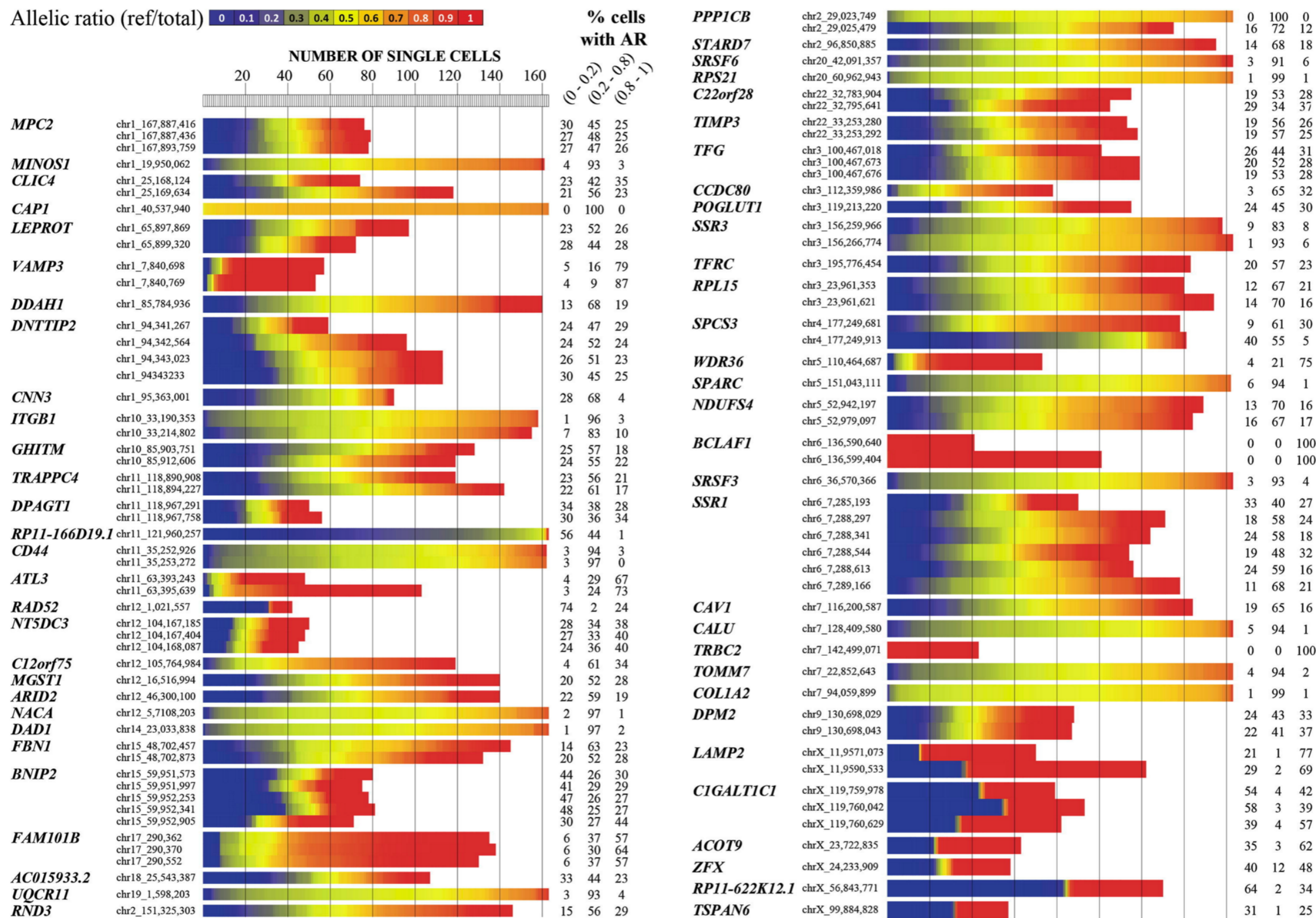




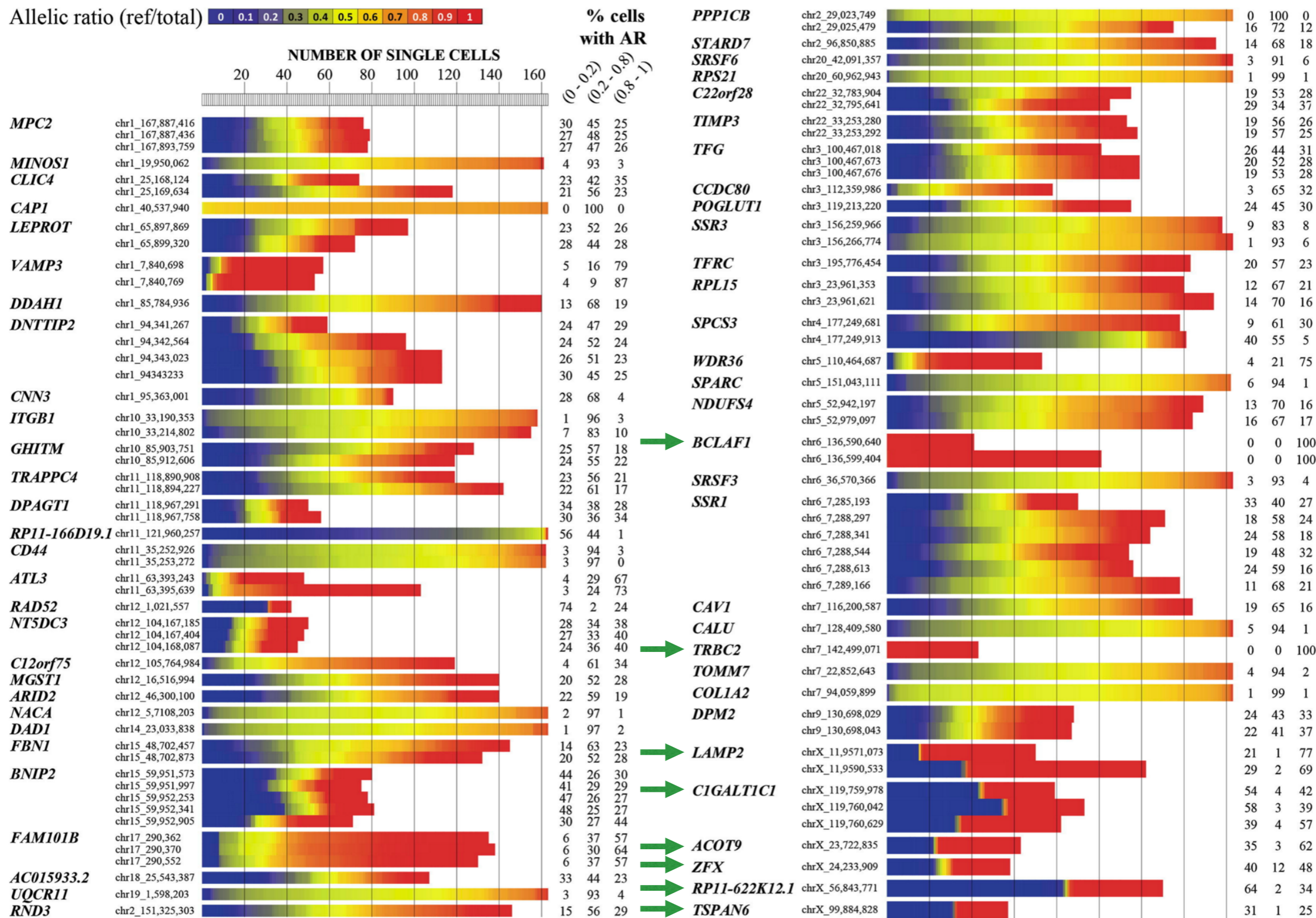
23.6 % of the hetSNVs had AR in 0.2 - 0.8 in UCF1014

34.2 % of the hetSNVs had AR in 0.2 - 0.8 in T2N

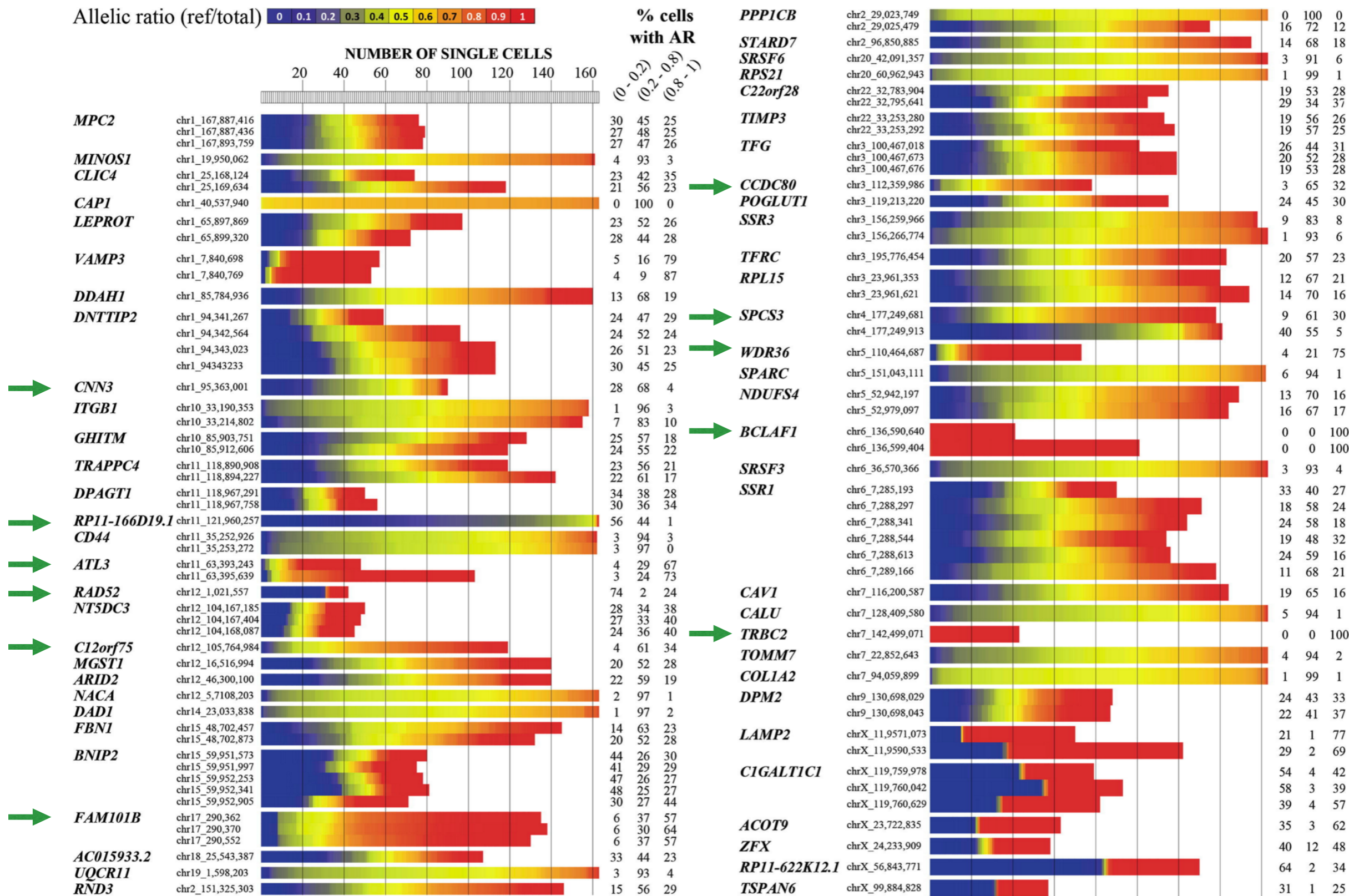




Representative subset (out of 568) genes w at least one hetSNP w RPSM >= 20 in >= 40 cells

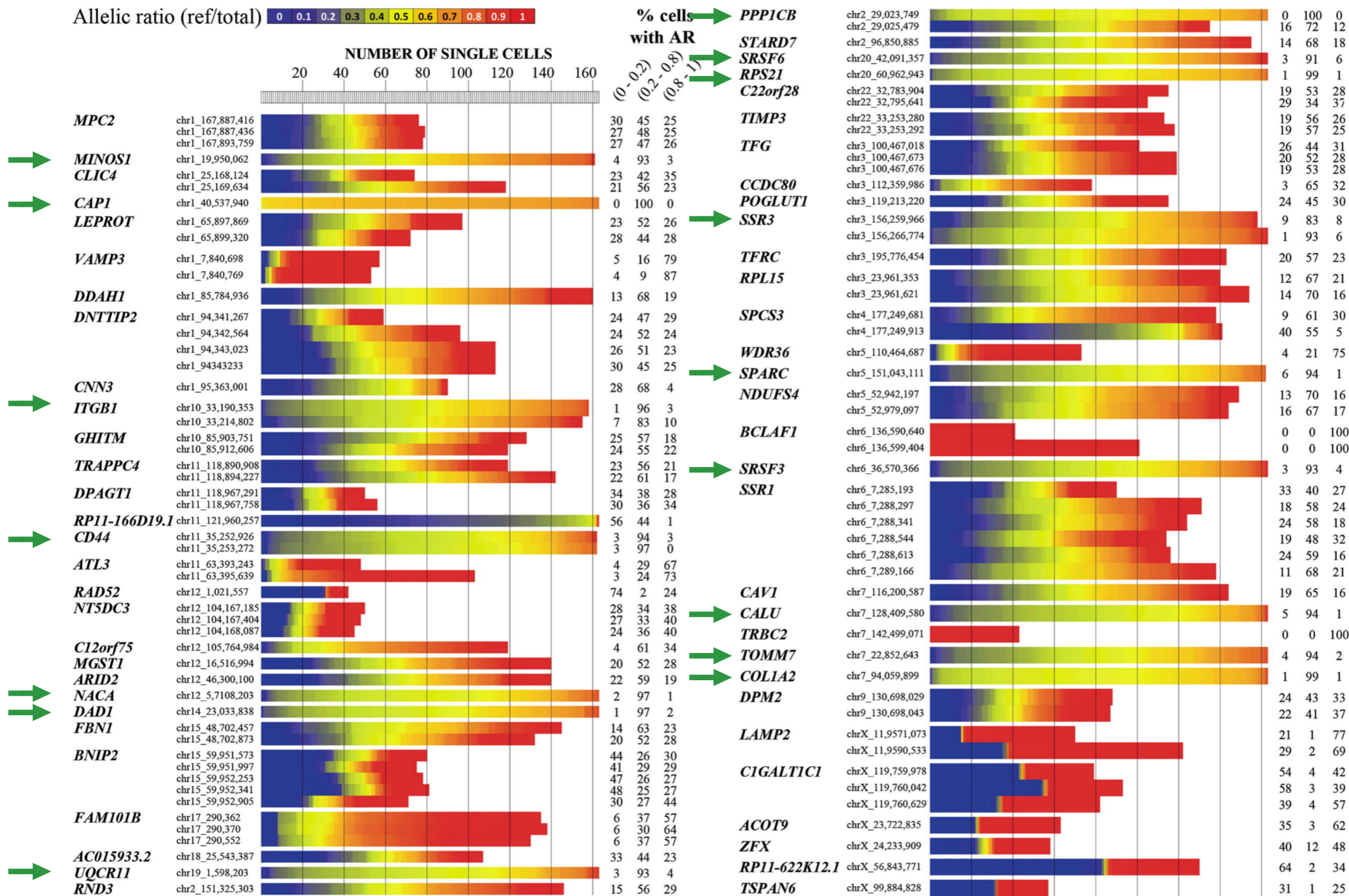


monoallelically expressed (<5% cells show biallelic expression)



skewed allelic expression (<10% of cell express one type of allele)





“single-cell biallelic” (>90% cells expressed both alleles, AR: 0.2-0.8)

Total transcript level, mRNA decay, initiation rate (averaged for genes in 163 sc)

