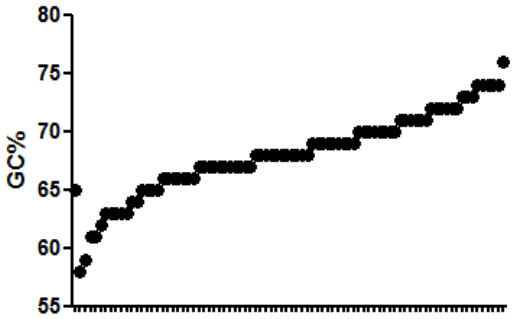
Supplementary Analysis (for other potential causes of missing genes)

GC contents

The GC% of 97 missing genes from the biological replicates of the one cell amplified samples were calculated using an online software, Oligo Calc (http://www.basic.northwestern.edu/biotools/oligocalc.html), and plotted. As shown in the graph below, the GC contents of all the missing genes examined are normally distributed, suggesting that there is no correlation between the GC% and amplification of these missing gene.

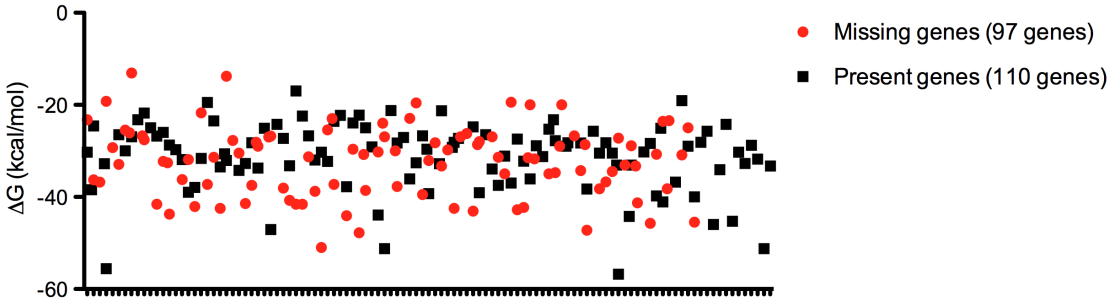


Transcript sizes

Among the 97 missing genes analyzed, only 43% are in computationally predicted operons based on the database of prokaryotic operons (http://www.Burkholderia.com/). More importantly, DNA random hexamers were used for reverse transcription. Therefore, the cDNA products should be shorter fragments, and transcript or operon sizes should not contribute to this bias.

RNA secondary structures

The secondary structure of RNA templates could potentially inhibit the reverse transcription (RT). To rule out this possibility, the Gibbs free energy values (ΔG) of these 100 missing genes were calculated using an online software, Zipfold (http://dinamelt.bioinfo.rpi.edu/zipfold.php), and plotted. As shown in the graph below, the ΔG values of all the missing genes examined overlap with the ones of 110 randomly chosen present genes, suggesting that there is no correlation between the RNA secondary structure and amplification of these missing gene. Furthermore, the fact that these ~100 missing genes were detected in the nonamplified sample, where RT was carried out identically, indirectly disproved this reason.



Gene expression diversity among individual cells in population

Arguably, the expression diversity among individual cells caused the "drop-out." A previous study (Taniguchi et al. 2010) showed that this diversity or "noise" exists for both high and low abundant transcripts. However, since we did not observe the diversity for the high abundant transcripts among individual cells (Fig. 5), then expression diversity among individual cells for low abundant transcripts is unlikely. Our data does

not agree with this previous study indicating that expression in individual cells was noisy (Taniguchi et al. 2010), and this discrepancy could be due to difference ways in which the single cells were treated and manipulated.

Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A, Xie XS. 2010. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**(5991): 533-538.