Published in final edited form as:

Wiley Interdiscip Rev RNA. 2017 November ; 8(6): . doi:10.1002/wrna.1433.

Single cell transcriptomics of non-coding RNAs and their cell-specificity

Katerina AB Gawronski and

Department of Genetics, University of Pennsylvania

Junhyong Kim

Department of Biology, Penn Program in Single Cell Biology, University of Pennsylvania

Abstract

Recent developments of single cell transcriptome profiling methods have led the realization that many seemingly homogeneous cells have surprising levels of expression variability. The biological implications of the high degree of variability is unclear but one possibility is that many genes are restricted in expression to small lineages of cells, suggesting the existence of many more cell types than previously estimated. Non-coding RNA (ncRNA) are thought to be key parts of gene regulatory processes and their single cell expression patterns may help dissect the biological function of single cell variability. Technology for measuring ncRNA in single cells is still in development and most of the current single cell datasets have reliable measurements for only lncRNA. Most works report that lncRNAs show lineage-specific restricted expression patterns, which suggests that they might determine, at least in part, lineage fates and cell subtypes. However, evidence is still inconclusive as to whether lncRNAs and other ncRNAs are more lineage-specific than protein coding genes. Nevertheless, measurement of ncRNAs in single cells will be important for studies of cell types and single cell function.

Keywords

single cell sequencing; RNA sequencing; noncoding RNA; long noncoding RNA; lncRNA; single-cell RNA-seq

Introduction

Metazoan development consists of the processes of morphogenesis and differentiation; that is, the spatial organization of cells into mechanical structures and the specialization of cells into particular functional phenotypes¹. The human body is estimated to be composed of 100 trillion cells, organized into molecularly distinct tissues and organs²⁻⁴ that arise through the complex process of differentiation. While the functional needs of different tissues and organs necessitate differentiation, morphologically and functionally similar cells within a given tissue were seen, until recently, as generally homogeneous building blocks. However, with increasingly higher resolution measurements it has become clear that individual

somatic cells display a wide range of molecular variation^{5–13}, including genomic and epigenomic variation^{14–16}. The ability to measure whole transcriptomes from single cells^{17, 18} has led to increasing numbers of putative cell types^{6, 19, 20} with significant cell-to-cell variation, even within clonal cell lines^{21, 22}.

The functional rationale for such broad molecular variation at the level of single cells is unclear. In Eberwine and Kim²³, we made the argument that, at least in part, the molecular states of any given cell should not have to meet exact stoichiometric requirements. That is, we proposed there might be multiple “equi-phenotypic” molecular states that can exist because different combinations of gene expression can lead to the same output²⁴. Thus, we proposed a “functionally neutral” view of single cell variation. On the other hand, there are many possible adaptive reasons for cell-to-cell variability including: bet-hedging against environmental perturbations^{10, 21}, combinatorial composition for tissue-level function²⁵, multi-cell ecological and signalling interactions^{12, 13, 22}, etc. (other possibilities reviewed in Dueck et al²⁶). The “functionally neutral” view and the “functionally adaptive” view of single cell variation are not mutually exclusive. It is possible that the variability in certain subsets of genes or patterns of covariation are functionally neutral, while variability in other gene subsets or patterns of covariation serves an adaptive purpose. Understanding functional versus neutral variation is a key problem in single cell biology, with significant implications for trans-differentiation, regenerative medicine, and tissue engineering. One possible approach to identifying functional single cell variability might be in focusing on the variability of non-coding RNAs, which tend to be predominantly involved in genomic regulatory function²⁷. In this review, we first briefly review non-coding RNAs and then examine the current evidence for single cell variation in non-coding RNA.

NON-CODING RNA

Although protein-coding sequences have thus far been the most highly studied sequences in the genome, non-coding sequences play crucial roles in a wide variety of cellular functions. Non-coding RNA is implicated in many different processes and pathways, including imprinting, differentiation, and cell cycle regulation. The classifications for non-coding RNA (ncRNA) are not universally agreed upon, but for this review we consider the following classes of ncRNA: tRNA, snoRNA or sdRNA, circRNA, miRNA, and lncRNA.

Transfer RNAs (tRNAs) are classic non-coding RNA molecules 70–100 nucleotides long that play a central role in the translation of mRNA into protein. Due to its crucial role in translation, mutations in mitochondrial tRNA and defects in the enzymes responsible for processing and posttranscriptional modification of tRNA are implicated in several human diseases including neurological disorders, cancer, and mitochondrial diseases^{28, 29}. Many of these diseases demonstrate heterogeneity in the degree to which different tissues are affected, which may be related to possible single cell variability. This heterogeneity is not unexpected, given that tRNA expression is closely linked to the translation needs of different tissues, and even varies between cells that are differentiating versus cells that are proliferating. Transfer RNAs can have complex genomic structure³⁰ and differentially regulated expression that affects protein translation rates through the availability of specific

anticodon carrying molecules³¹, which potentially suggests they may also play a role in cell-to-cell variability.

Small nucleolar RNAs (snoRNAs) are ncRNAs of 60–300 bp that are primarily involved in modifying and processing ribosomal RNA³². snoRNAs are subunits of small nucleolar ribonucleoproteins, which post-transcriptionally modify ribosomal RNA through 2'-O-methylation and pseudouridylation. Ribosomes play a central role in maintaining proper cell function through protein synthesis, and snoRNA expression has been correlated with the proliferation of tumor cells, and genes containing intronic snoRNA have been linked to several cancers. The loss of several snoRNAs has also been associated with neurological disorders such as Prader-Willi syndrome. Changes in snoRNA expression on a cellular level have been shown to result in oxidative stress phenotypes³³. Of note, some results suggest that snoRNA interacts with long non-coding RNA to influence alternative splicing³⁴.

Circular RNAs (circRNAs) are recently described closed, circular RNA molecules that arise through the back-splicing of precursor mRNA exons³⁵. Several studies indicate that more than 10% of expressed genes are capable of producing circRNAs. The function of circRNAs as a class is not well understood, although recent research indicates that several circRNAs are involved in gene regulation. For example, circRNAs can act as sponges for miRNA, as well as impact alternative splicing of the transcripts that produce circRNAs. Some circRNAs also demonstrate cell-type and tissue-type specificity³⁶.

Micro RNAs (miRNA) are single-stranded non-coding RNAs that are approximately 22 bp in length broadly involved in both transcriptional and translational regulation^{37, 38}. miRNAs are formed when long pri-miRNA transcripts are processed first by protein Drosha to form hairpin pre-miRNA. This pre-miRNA is then cleaved by Dicer, resulting in a double-stranded molecule. One of the strands of this molecule is then incorporated into Argonaute to form the RNA-induced silencing complex, which binds to mRNA sequences that base-pair with the single-stranded miRNA molecule to either inhibit translation or induce decay of the bound mRNA. Given its mechanism of action, the amount of miRNA in a cell dictates the impact it has on gene expression, and therefore single cell levels of miRNA are of particular interest. miRNA plays an important role in development and differentiation, as well as in a variety of cancers and cardiovascular diseases^{39, 40}.

Long non-coding RNAs (lncRNAs) are RNA molecules of 200 bp or longer that do not code for protein. As cataloged by ENCODE (v21), there are at least 15,877 known lncRNAs to date⁴¹. LncRNAs perform a wide range of functions, including imprinting, translation, splicing, cell differentiation, and cell cycle control. Given their extensive involvement in the regulation of gene expression and the cell cycle, lncRNAs play a key role in development and have been associated with a variety of human diseases including cancer, neurodegenerative diseases, and cardiovascular disease. LncRNAs are known to demonstrate tissue-specific and cell-type specific expression patterns, and research suggests that tissue-specific lncRNAs play a role in disease^{42, 43}. A subcategory of lncRNAs of particular interest is long intergenic non-coding RNAs (lincRNAs), which are thought to comprise the majority of lncRNAs. Unlike some long non-coding RNAs, long intergenic non-coding RNAs do not overlap any exons or transcripts of known genes and are instead transcribed

from intergenic regions. LincRNAs play key roles in pluripotency and differentiation, by influencing global gene expression and repressing lineage programs⁴⁴.

SINGLE CELL MEASUREMENTS OF NON-CODING RNA

Until recently, ncRNA expression has been measured by averaging transcriptomes of bulk RNA from many different cells (e.g.,^{45, 46}). Many non-coding RNAs are expressed at low levels, are transiently expressed, or are coupled to other regulatory transcription events. Because of these characteristics, bulk measurements limit sensitivity to detect non-coding RNA expression. Obtaining expression measurements from heterogeneous pools of cells also makes it difficult to interrogate the key roles of ncRNA in differentiation, where a small subset of cells might initiate substantial, wide-reaching changes in the overall genomic program. To evaluate the role of ncRNAs in individual cells, recent studies have turned to single cell transcriptome sequencing. One cannot assay the whole transcriptome of a single cell (typically with ~10pg of total RNA and ~0.2 pg of mRNA), without RNA amplification. The two major methods of single cell RNA amplification, in vitro transcription methods⁴⁷ and PCR-based methods⁴⁸, both utilize 3' poly-A tails to prime at least one direction of amplification. Poly-A priming also has the advantage of excluding the rRNA that comprises the vast majority of RNA. While certain non-coding RNAs such as pri-miRNAs or poly-A tailed lncRNAs can be recovered with standard single cell RNA sequencing, other ncRNAs such as mature miRNAs cannot be assayed by poly-A dependent methods. Recently, several methods have been developed to specifically target small RNAs (such as miRNA) from single cells. Faridani et al.⁴⁹ use adaptor ligations to both 5' and 3' ends while using masking oligonucleotides to eliminate ribosomal RNAs. The authors report capturing 3,800 miRNA, 3,500 tsRNA (tRNA-derived small RNA), and 600 sdRNA (sno-RNA derived RNA) per cell, a surprisingly large number of small non-coding RNAs with likely regulatory effects. The authors report that this comprehensive profile of single cell miRNAs was able to robustly separate sub-classes of cells. These results are in contrast to the results of studies using standard single cell RNA sequencing methods (pri-miRNA)—many of these studies have failed to find cell-specific expression patterns of miRNAs⁵⁰. These differing findings suggest that small ncRNAs may have distinct single cell signatures, but that these signatures cannot be adequately captured using the usual single cell RNA sequencing approaches. Randomly anchored primers have also been used to amplify non-poly-A RNAs from single cells⁵¹. Using this method, Fan et al. detected unique circular RNAs both from HEK293T cells and also in early stages of developing mouse embryos. Fine-scale high-resolution assays of single cells seem to identify a high degree of novel circular RNA expression, which may help dissect the function of these cryptic ncRNA. Both Faridani et al.⁴⁹ and Fan et al.⁵¹ report good performance of their amplification methods compared to control standards. However, adaptor ligation is an inefficient process, and the total recovery and expression representativeness of non-poly-A amplification methods remain to be validated. Most of the current literature on single cell ncRNA report on lncRNAs, whose expression can be mostly recovered by standard poly-A anchored methods – therefore, we will focus our discussion on patterns of lncRNA expression.

SINGLE CELL DETECTION AND VARIATION OF LNCRNAS

Single cell transcriptome assays are subject to many kinds of technical problems, and the resulting data must be corrected for bias and analyzed with the appropriate statistical analyses⁵². Nevertheless, one of the key advantages of single cell transcriptome analysis is that these assays can detect rare transcripts that are present in only a minority of cells⁵³; thus, many single cell studies discover novel non-coding transcripts. Re-analyzing public single cell lncRNA data, Zhang et al.⁵⁴ report finding 5,563 novel lncRNAs expressed in mouse early embryos. Using single cell assays of CD133+/GFAP-ependymal cells from the mouse adult subventricular zone, Luo et al.⁵⁵ report 600 novel multi-exonic lncRNA that had not been previously reported from this region, even in microdissected material. Such novel RNAs are often not revealed even in deep bulk sequencing, and certain lncRNAs are reported to be highly cell type specific, such as LOC646329 lncRNA found to be expressed only in radial glial cells of the human neocortex^{56, 57}. Similarly, novel expression of previously unknown lncRNAs was identified in single cell studies of hematopoietic stem cell formation⁵⁸. More recently, some groups have been carrying out single-nuclear sequencing^{59, 60}. For many single cells, intact recovery of the total transcriptome cannot be achieved due to the difficulty of capturing delicate cells or dissociating the cells from tissue slices. Often, the nucleus is more efficiently collected, and therefore nuclear sequencing represents a possible robust high-throughput path for large-scale single cell analysis. The nuclear fraction of the transcriptome is enriched in non-coding RNAs, especially lncRNAs. Previously, many studies examining whole cell transcriptomes did not find a strong relationship between the expression of various ncRNA and mRNA, but studies that conducted single-nucleus sequencing report stronger correlated patterns of expression for lncRNA, pri-miRNA, and mRNA⁶¹. In particular, this study⁶¹ suggests that the differences in lncRNA expression between cell soma and cell nucleus are especially cell-specific, possibly because nuclear RNA expression more closely represents the dynamics of gene regulation.

One of the most interesting observations from the emerging body of single cell transcriptome analyses has been the extent of cell-to-cell variation in mRNA expression. Such heterogeneity also seems to hold true for ncRNA expression. A study of ~2,000 lncRNAs in 380+ cells from both primary glioblastoma and glioblastoma stem-like cell lines found significant levels of cell-to-cell variation of lncRNA, even within the established clonal cell lines⁶². In addition, the authors found significant splice pattern variation, suggesting that single cell variation of lncRNAs manifests in both sequence and expression. Multiple splice variants were expressed simultaneously in many cells but displayed variation in their relative expression levels from cell-to-cell. Long non-coding RNAs have also demonstrated single cell variability in other cancers. For example, the novel antisense lncRNA HIPSTR was found when researchers deep sequenced prostate cancer cell lines⁶³. This novel lncRNA was found to be expressed during the major wave of embryonic genome activation in the 8-cell stage. In general, the study suggested that lncRNAs are more than twice as likely to have significant cell-to-cell variability as mRNAs. It is not clear whether this is due to the technical difficulty of measuring lowly expressed genes, (such as lncRNAs), but overall these studies suggest that lncRNAs are at least as variable in their expression patterns as protein-coding genes.

Interestingly, single cell variation of the transcriptome might itself be regulated by non-coding RNA. Gambardella et al.⁶⁴ examined single cell transcriptome variation after the introduction of two miRNAs into miRNA-deficient DGCR8 knockout mice. They found that miRNA let-7c increased single cell transcriptome variability while miRNA miR-294 decreased single cell variability. The effects on cell transcriptome variability were especially pronounced for pluripotency regulators. A key finding here was that the increase in heterogeneity by let-7c seemed to arise from the generation of distinct sub-lineages of cell types, rather than a simple increase in stochastic gene expression. The results of this study are consistent with the idea that much of single cell variability is due to the generation of cell sub-types, and that these sub-types are either associated with or regulated by non-coding RNAs.

Does single cell lncRNA expression heterogeneity reflect hidden sub-types of cells?

As mentioned above, increasing numbers and types of non-coding RNAs have been identified through recent RNA sequencing efforts, and many of these molecules have been speculated to play a role in gene and cell phenotype regulation. This is particularly the case for lncRNAs, which have long been associated with developmental processes. If lncRNAs are in fact major regulatory factors of cell differentiation and cell fates, we would expect them to show restricted expression in small numbers of single cells, support distinct clustering of cell types, and show statistical signatures of high variability.

In one of the earliest single cell studies of human development, Yan et al.⁶⁵ studied 124 cells from donor preimplantation embryos and found approximately 18,000 lncRNAs expressed in these early developmental stages (out of 28,640 known at the time). That is, in these key early stages more than 60% of all known lncRNAs were expressed. Single cell RNA-seq's higher resolution also led to the annotation of 2,733 novel lncRNAs in these embryonic cells. When using bulk sequencing methods, lncRNAs were thought to be expressed at ~10% of the expression level of protein-coding mRNAs. However, at the single cell level, expression of lncRNA jumps to ~40% of the expression level of protein-coding genes, suggesting that restriction of lncRNA expression to specific cell lineages is the cause of its overall low expression in bulk measurements. The study's statistical analysis suggests that lncRNAs were specifically regulated during the studied developmental stages and that their expression was not due to random leaky expression. In the human neocortex, Liu et al.⁵⁶ also found many lncRNAs that were abundantly expressed in single cells. These lncRNAs generally had low expression in bulk RNA sequencing data, suggesting their expression occurs in a minority of single cells. Again, the data supports the idea that for many cell-specific genes, low expression in bulk RNA-seq data corresponds not to low expression in individual cells, but to rare expression within a population of cells. In Liu et al.⁵⁶, the authors report that lncRNAs with low abundance in bulk measurements tend to be more cell-specific at the single cell level and that a statistical cell-specificity score of lncRNAs is comparable to those of mRNAs. The authors also tested the functional consequence of one cell-specific lncRNA, LOC646329, by CRISPRi knockdown and found that the loss of this lncRNA reduced cell proliferation in a particular sub-population of cells. This reduced proliferation as a result of lncRNA knockdown suggests that lncRNAs play a critical role in determining the fate of certain cell lineages. Qui et al.⁵⁷ reanalyzed public single cell

embryonic developmental data from human and mouse and reported that human lncRNAs tend to have developmental stage-specific expression, while the mouse expression profile seems to follow a more continuous time-dependent trajectory. In particular, lncRNAs are reported to have more time-dependent expression signatures than protein-coding genes. Interestingly, the authors suggest that the expression of lncRNAs involved in maternal to zygotic expression transition seems to be conserved between human and mouse, while those at other developmental stages did not show conserved expression patterns. In Johnson et al.⁶⁶ the authors dissociated and fractionated cells from human fetal brain tissue and identified many lncRNAs specifically enriched in apical radial glial cells while displaying a smaller and distinct pattern for outer radial glial (ORG) cells. As with Qui et al.⁵⁷ study, cell-specific lncRNAs seem to be comprised of less conserved, human-specific lncRNAs that lack any mouse homologs. Interestingly, the authors report that many of the ORG-specific lncRNAs are conserved in non-rodent non-primate species in a pattern of parallel evolution (or conserved evolution with a derived pattern in rodents). Further, these lncRNAs seem to be associated with gyrification of the brains. This suggests that lncRNAs that are associated with minor sub-types of cells might also be associated with specific developmental processes involved in morphogenesis.

Measurements of single cells during cell fate reprogramming also suggest important lineage-specific roles for lncRNA. Kim et al.⁶⁷ mapped 437 lncRNAs expressed at significant levels (greater than 10 RPKM) for iPS and ES cells and found dynamic expression changes in transitional cells. They suggest that polycomb-bound lncRNAs may repress developmental genes in a heterogeneous manner across individual cells. The authors suggest support for lineage-specific expression of lncRNA, which may partly explain the fact that many lncRNAs show significant cell-to-cell variability. Additionally supporting the hypothesis that lncRNAs are involved in single cell lineage-specific expression, Petropoulos et al.⁶⁸ found variability of sex-specific cell lineage expression for XIST and XACT (X-linked lncRNA covering XIST free X chromosomes), and this variability was modulated across developmental stages.

Cell-specific expression of lncRNAs has also been suggested to manifest during tumor formation. In a direct study of 500+ cells from five glioblastomas and two glioblastoma cell lines, the authors found that more than 1,000 lncRNAs significantly varied between at least two cells within the MGH29 glioblastoma cell line. Further evidence of lncRNA expression heterogeneity was observed during tumor progression, where the authors report frequent gain/loss of lncRNA expression⁶⁹. Interestingly, while lncRNA expression was variable across individual cells, the expression patterns were more correlated within each of the five samples compared to across the tumor samples, suggesting a tissue-of-origin effect to specific lncRNA expression. In fact, of the 500 least variable lncRNAs, only 175 were shared across the five samples. Consistent with previous studies, evolutionarily conserved lncRNAs were less variable across the samples, although a higher fraction of conserved lncRNAs was observed in single cells compared to non-conserved human-specific lncRNAs. Focusing on long intervening non-coding RNAs (lincRNA) in 37 HeLa-S3 cells, Wang et al.⁷⁰ found an average of 511 lincRNAs expressed in each cell and observed expression of more than 2,000 total lincRNAs. In the same study, the authors found that ~68% of lincRNAs were expressed in fewer than ten cells, compared to ~34% of protein-coding

genes. Further, lincRNAs displayed significantly higher levels of cell-specificity than protein-coding genes. Also re-examining published single cell RNA-seq data from human pre-implantation embryos⁶⁵, Lv et al.⁷¹ found more than 4,400+ novel lincRNAs and observed significantly greater tissue-specificity for the lincRNAs compared to protein-coding genes. In fact, the novel lincRNAs tended to show higher tissue-specificity than previously known lincRNAs. In sum, many of the single cell RNA-seq studies of lincRNA and lincRNA find low levels of expression explained by the restriction of expression to a minority of cells, perhaps indicating their role in differentiating particular cell lineages.

Utilizing a method different from single cell RNA-seq, Cabili et al.⁷² used single molecule FISH for 61 different lincRNA probes to assay single cell expression patterns in human fibroblasts (foreskin and lung) and HeLa cells. They found distinct and heterogeneous patterns of sub-cellular lincRNA localizations, with the majority (55%) of the lincRNAs found predominantly in the nucleus. However, distinct from the papers above, they did not find that lowly expressed lincRNAs found in bulk sequencing corresponded to expression in minority of cells. Most of the lincRNAs were found to have similar localization patterns across the different cell types, with only five lincRNAs displaying possible cell-type-specific localization patterns. However, the authors attribute the results to low expression levels and possibly spurious localization patterns. In general, they found cell-to-cell variability of lincRNA at levels similar to protein-coding genes, but they did not find evidence for a small number of cells with high lincRNA expression, which might be expected for lineage-dependent restricted expression of lincRNAs.

Conclusion

Some of the earliest single cell datasets such as Shalek et al.⁷³ reported bimodality in single cell variation, especially for lincRNAs, and suggested cell-specific expression in a minority of the cells as an explanation for the pattern. The majority of the papers that examined single cell RNA-seq data sets suggest cell-specific restriction of lincRNA expression, which might also indicate that they play important roles in differentiating minor lineages of the cells. Evolutionary analyses also hint that these cell-specific patterns are especially pronounced for recent species-specific (e.g., human-specific) lincRNAs, which is consistent with the idea that these cell-specific lincRNAs evolved to direct the differentiation of species-specific cell fates. Yet, the orthogonal methodological study of Cabili et al.⁷² using single molecule FISH shows no indication that lincRNAs are any more lineage-specific than protein-coding genes. On the one hand, single cell RNA-seq studies are still few in number, and many of the studies are re-analyses of previously collected datasets. On the other hand, the Cabili et al.⁷² study only examined a handful of lincRNA loci (61 probes) out of the thousands of lincRNA reported by other studies. Whether lincRNAs are lineage-restricted and play an important role in establishing such lineage fates is still unresolved and will require additional data. In particular, Cabili et al. kind of analysis will need to be redone with a larger selection of probes including species-specific lincRNA probes.

Given the preponderance of evidence from recent single cell papers, it is clear that individual cells, even clonal cells, vary considerably in their transcriptome states. Non-coding RNAs seem to show at least as much single cell variability as mRNA, and many studies suggest

that certain classes of non-coding RNA may be even more variable, especially regarding the establishment of new sub-populations of cell types. Cell lineage fate determination may be a major reason for single cell variability. Even after developmental differentiation, we hypothesize that there may be many functional rationales for single cell variation²⁶. In particular, one hypothesis is that cells must be inherently heterogeneous to allow for graded responses or dynamics. For example, every skin cell turns over multiple times in a year, but the turnover is not synchronous—some degree of molecular heterogeneity governs slow and rare turnover of individual cells. In a careful and fascinating study Wu et al., examined the transcriptional dynamics underlying B cell class switching⁷⁴. Non-coding transcripts upstream of immunoglobulin constant regions are required for B cell class switching (IgG1 and IgE). Using single cell RNA-seq and single cell fluorescent reporters, the authors found that differential activation thresholds exist for two promoters related to the class switching. Further, they found that the heterogeneity in promoter activation was characteristic of this pathway, even in synchronized highly homogeneous cells. They find that the two promoters each have a well-defined probability of activation with inherent stochasticity that is structured to produce a system level cell-to-cell variation and cell type switching decisions. Thus, the study of these loci suggests a specific molecular mechanism of non-coding RNA and promoters that shape the cell-to-cell heterogeneity distribution.

Noncoding RNA are known to exert their effects through a variety of mechanisms including chromatin remodeling, serving as protein scaffolds, binding to and inhibiting the translation of mRNA, binding to and changing the splicing patterns of mRNA transcripts, and acting as sponges for miRNA^{75, 76}. While the functions of some specific ncRNA are well-characterized, the mechanisms of the majority of ncRNA are unknown, and it is yet to be determined whether these mechanisms have single-cell-specific effects. The studies discussed in this review provide some insight into how specific ncRNAs may exert their effects, rather than insights into novel mechanisms for entire classes of ncRNA at the single cell level. For example, Liu et. al 2016 evaluated the role of *LOC646329* in single radial glial cells, whose knockdown seems to impact cell proliferation⁵⁶. Qui et al. 2016 found that lncRNA and gene co-expression modules at the 8-cell stage were enriched in functions related to nucleosome and chromatin assembly, suggesting that lncRNAs may regulate nucleosome and chromatin assembly during zygotic genome activation⁵⁷. Gambardella et al. 2016 found that miR-294 expression was associated with homogenous transcriptomes across single cells, while let-7 was associated with transcriptional heterogeneity across single cells as a result of the creation of cell subpopulations⁶⁴. The location of ncRNA within a cell may further help narrow the scope of its potential mechanism of action, especially if they differ in different cells. Future studies that combine single cell sequencing, knockdown experiments, and localization of ncRNA within single cells may be able to gain further insight into ncRNA function, and whether these functions vary at the level of single cells.

If non-coding RNAs are important markers of cell lineages and determinants of phenotypes, they might be an important focus for future systematic studies of cell types. But more importantly, mammalian tissues, organs, and individuals are comprised of trillions of individual cells, which in aggregate must carry out quantitative functions. The output from each single cell comprises a quantum unit that sum over all the cells to produce the tissue level function. The way in which single cell variability is structured determines the sum over

the quantum units and the quantitative characteristics of tissue function. Future studies of single cell variation, especially coupled to functional outputs, may finally allow a more mechanistic understanding of how development and normal tissue function coordinates functional responses with quantitative precision.

Acknowledgments

We thank Hannah Dueck for early assistance with this manuscript. Hannah helped curate many of the papers discussed in this paper. This work has been funded in part by NIMH grant U01MH098953 to J.K and J. Eberwine and GM008216 T32 grant to K.G. This work is also supported in part by a Health Research Formula Fund from the Commonwealth of Pennsylvania. The funding agencies had no role in the results or opinions of this paper.

References

1. Waddington, CH. Principles of Embryology. London, UK: George Allen & Unwin; 1956.
2. Consortium GT. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. 2015; 348:648–660. [PubMed: 25954001]
3. Grundberg E, Small KS, Hedman AK, Nica AC, Buil A, Keildson S, Bell JT, Yang TP, Meduri E, Barrett A, et al. Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet*. 2012; 44:1084–1089. [PubMed: 22941192]
4. Mele M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, Young TR, Goldmann JM, Pervouchine DD, Sullivan TJ, et al. Human genomics. The human transcriptome across tissues and individuals. *Science*. 2015; 348:660–665. [PubMed: 25954002]
5. Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science*. 2015:348.
6. Chiu IM, Barrett LB, Williams EK, Strohlic DE, Lee S, Weyer AD, Lou S, Bryman G, Roberson DP, Ghasemlou N, et al. Transcriptional profiling at whole population and single cell levels reveals somatosensory neuron molecular diversity. *eLife*. 2014
7. Deng Q, Ramskold D, Reinius B, Sandberg R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science (New York, NY)*. 2014; 343:193–196.
8. Dueck H, Khaladkar M, Kim TK, Spaethling JM, Francis C, Suresh S, Fisher SA, Seale P, Beck SG, Bartfai T, et al. Deep sequencing reveals cell-type-specific patterns of single-cell transcriptome variation. *Genome biology*. 2015:16. [PubMed: 25622821]
9. Islam S, Kjällquist U, Moliner A, Zajac P, Fan J-B, Lönnerberg P, Linnarsson S. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome research*. 2011; 21:1160–1167. [PubMed: 21543516]
10. Lee M-CW, Lopez-Diaz FJ, Khan SY, Tariq MA, Dayn Y, Vaske CJ, Radenbaugh AJ, Kim HJ, Emerson BM, Pourmand N. Single-cell analyses of transcriptional heterogeneity during drug tolerance transition in cancer cells by RNA sequencing. *Proceedings of the National Academy of Sciences*. 2014; 111:E4726–E4735.
11. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, Cahill DP, Nahed BV, Curry WT, Martuza RL, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science*. 2014; 344:1396–1401. [PubMed: 24925914]
12. Shalek AK, Satija R, Shuga J, Trombetta JJ, Gennert D, Lu D, Chen P, Gertner RS, Gaublotte JT, Yosef N, et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature*. 2014; 510:363–369. [PubMed: 24919153]
13. Tanay A, Regev A. Scaling single-cell genomics from phenomenology to mechanism. *Nature*. 2017; 541:331–338. [PubMed: 28102262]
14. Erwin JA, Marchetto MC, Gage FH. Mobile DNA elements in the generation of diversity and complexity in the brain. *Nat Rev Neurosci*. 2014; 15:497–506. [PubMed: 25005482]
15. Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature*. 2005; 435:903–910. [PubMed: 15959507]

16. Smallwood SA, Lee HJ, Angermueller C, Krueger F, Saadeh H, Peet J, Andrews SR, Stegle O, Reik W, Kelsey G. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nature Methods*. 2014; 11:817–820. [PubMed: 25042786]
17. Eberwine J, Sul JY, Bartfai T, Kim J. The promise of single-cell sequencing. *Nat Methods*. 2014; 11:25–27. [PubMed: 24524134]
18. Sul JY, Wu CwK, Zeng F, Jochems J, Lee MT, Kim TK, Peritz T, Buckley P, Cappelleri DJ, Maronski M, et al. Transcriptome transfer produces a predictable cellular phenotype. *Proceedings of the National Academy of Sciences*. 2009; 106:7624–7629.
19. Shekhar K, Lapan SW, Whitney IE, Tran NM, Macosko EZ, Kowalczyk M, Adiconis X, Levin JZ, Nemesh J, Goldman M, et al. Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell*. 2016; 166:1308–1323. e1330. [PubMed: 27565351]
20. Usoskin D, Furlan A, Islam S, Abdo H, Lonnerberg P, Lou D, Hjerling-Leffler J, Haeggstrom J, Kharchenko O, Kharchenko PV, et al. Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nat Neurosci*. 2015; 18:145–153. [PubMed: 25420068]
21. Kim K-T, Lee HW, Lee H-O, Kim SC, Seo YJ, Chung W, Eum HH, Nam D-H, Kim J, Joo KM, et al. Single-cell mRNA sequencing identifies subclonal heterogeneity in anti-cancer drug responses of lung adenocarcinoma cells. *Genome biology*. 2015;16. [PubMed: 25622821]
22. Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. 2016; 352:189–196. [PubMed: 27124452]
23. Eberwine J, Kim J. Cellular Deconstruction: Finding Meaning in Individual Cell Variation. *Transactions on Computational Biology and Bioinformatics*.
24. Kim J, Eberwine J. RNA: state memory and mediator of cellular phenotype. *Trends in cell biology*. 2010; 20:311–318. [PubMed: 20382532]
25. Marder E, Goaillard J-M. Variability, compensation and homeostasis in neuron and network function. *Nature Reviews Neuroscience*. 2006; 7:563–574. [PubMed: 16791145]
26. Dueck H, Eberwine J, Kim J. Variation is function: Are single cell differences functionally important?: Testing the hypothesis that single cell variation is required for aggregate function. *Bioessays*. 2016; 38:172–180. [PubMed: 26625861]
27. Morris KV, Mattick JS. The rise of regulatory RNA. *Nature reviews Genetics*. 2014; 15:423–437.
28. Grewal SS. Why should cancer biologists care about tRNAs? tRNA synthesis, mRNA translation and the control of growth. *Biochim Biophys Acta*. 2015; 1849:898–907. [PubMed: 25497380]
29. Torres AG, Batlle E, Ribas de Pouplana L. Role of tRNA modifications in human diseases. *Trends Mol Med*. 2014; 20:306–314. [PubMed: 24581449]
30. Randau L, Soll D. Transfer RNA genes in pieces. *EMBO Rep*. 2008; 9:623–628. [PubMed: 18552771]
31. Kirchner S, Ignatova Z. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nat Rev Genet*. 2015; 16:98–112. [PubMed: 25534324]
32. Bachellerie JP, Cavaille J, Huttenhofer A. The expanding snoRNA world. *Biochimie*. 2002; 84:775–790. [PubMed: 12457565]
33. Stepanov GA, Filippova JA, Komissarov AB, Kuligina EV, Richter VA, Semenov DV. Regulatory role of small nucleolar RNAs in human diseases. *Biomed Res Int*. 2015; 2015:206849. [PubMed: 26060813]
34. Li R, Fox AH. SPArking Interest in the Long Noncoding RNA World: A New Class of 5' SnoRNA-Stabilized LncRNA that Influences Alternative Splicing. *Mol Cell*. 2016; 64:435–437. [PubMed: 27814487]
35. Hentze MW, Preiss T. Circular RNAs: splicing's enigma variations. *EMBO J*. 2013; 32:923–925. [PubMed: 23463100]
36. Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S. circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell*. 2014; 56:55–66. [PubMed: 25242144]
37. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell*. 2009; 136:215–233. [PubMed: 19167326]

38. Huang Y, Shen XJ, Zou Q, Wang SP, Tang SM, Zhang GZ. Biological functions of microRNAs: a review. *J Physiol Biochem.* 2011; 67:129–139. [PubMed: 20981514]
39. Li Y, Kowdley KV. MicroRNAs in common human diseases. *Genomics Proteomics Bioinformatics.* 2012; 10:246–253. [PubMed: 23200134]
40. Mendell JT, Olson EN. MicroRNAs in stress signaling and human disease. *Cell.* 2012; 148:1172–1187. [PubMed: 22424228]
41. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012; 489:57–74. [PubMed: 22955616]
42. Lorenzen JM, Thum T. Long noncoding RNAs in kidney and cardiovascular diseases. *Nat Rev Nephrol.* 2016; 12:360–373. [PubMed: 27140855]
43. Wapinski O, Chang HY. Long noncoding RNAs and human disease. *Trends Cell Biol.* 2011; 21:354–361. [PubMed: 21550244]
44. Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, Young G, Lucas AB, Ach R, Bruhn L, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature.* 2011; 477:295–300. [PubMed: 21874018]
45. Wang ET, Sandberg R, Luo S, Khrebukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. *Nature.* 2008; 456:470–476. [PubMed: 18978772]
46. Weber JA, Baxter DH, Zhang S, Huang DY, Huang KH, Lee MJ, Galas DJ, Wang K. The microRNA spectrum in 12 body fluids. *Clin Chem.* 2010; 56:1733–1741. [PubMed: 20847327]
47. Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, Eberwine JH. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc Natl Acad Sci U S A.* 1990; 87:1663–1667. [PubMed: 1689846]
48. Ramsköld D, Luo S, Wang Y-C, Li R, Deng Q, Faridani OR, Daniels GA, Khrebukova I, Loring JF, Laurent LC, et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nature biotechnology.* 2012; 30:777–782.
49. Faridani OR, Abdullayev I, Hagemann-Jensen M, Schell JP, Lanner F, Sandberg R. Single-cell sequencing of the small-RNA transcriptome. *Nat Biotechnol.* 2016; 34:1264–1266. [PubMed: 27798564]
50. Spaethling JM, Na YJ, Lee J, Ulyanova AV, Baltuch GH, Bell TJ, Brem S, Chen HI, Dueck H, Fisher SA, et al. Primary Cell Culture of Live Neurosurgically Resected Aged Adult Human Brain Cells and Single Cell Transcriptomics. *Cell Rep.* 2017; 18:791–803. [PubMed: 28099855]
51. Fan X, Zhang X, Wu X, Guo H, Hu Y, Tang F, Huang Y. Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome Biol.* 2015; 16:148. [PubMed: 26201400]
52. Dueck HR, Ai R, Camarena A, Ding B, Dominguez R, Evgrafov OV, Fan JB, Fisher SA, Herstein JS, Kim TK, et al. Assessing characteristics of RNA amplification methods for single cell RNA sequencing. *BMC Genomics.* 2016; 17:966. [PubMed: 27881084]
53. Eberwine J, Lovatt D, Buckley P, Dueck H, Francis C, Kim TK, Lee J, Lee M, Miyashiro K, Morris J, et al. Quantitative biology of single neurons. *J R Soc Interface.* 2012; 9:3165–3183. [PubMed: 22915636]
54. Zhang K, Huang K, Luo Y, Li S. Identification and functional analysis of long non-coding RNAs in mouse cleavage stage embryonic development based on single cell transcriptome data. *BMC Genomics.* 2014; 15:845. [PubMed: 25277336]
55. Luo Y, Coskun V, Liang A, Yu J, Cheng L, Ge W, Shi Z, Zhang K, Li C, Cui Y, et al. Single-cell transcriptome analyses reveal signals to activate dormant neural stem cells. *Cell.* 2015; 161:1175–1186. [PubMed: 26000486]
56. Liu SJ, Nowakowski TJ, Pollen AA, Lui JH, Horlbeck MA, Attenello FJ, He D, Weissman JS, Kriegstein AR, Diaz AA, et al. Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biol.* 2016; 17:67. [PubMed: 27081004]
57. Qiu JJ, Ren ZR, Yan JB. Identification and functional analysis of long non-coding RNAs in human and mouse early embryos based on single-cell transcriptome data. *Oncotarget.* 2016; 7:61215–61228. [PubMed: 27542205]

58. Zhou F, Li X, Wang W, Zhu P, Zhou J, He W, Ding M, Xiong F, Zheng X, Li Z, et al. Tracing haematopoietic stem cell formation at single-cell resolution. *Nature*. 2016; 533:487–492. [PubMed: 27225119]
59. Habib N, Li Y, Heidenreich M, Swiech L, Avraham-Davidi I, Trombetta JJ, Hession C, Zhang F, Regev A. Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science*. 2016; 353:925–928. [PubMed: 27471252]
60. Lake BB, Ai R, Kaeser GE, Salathia NS, Yung YC, Liu R, Wildberg A, Gao D, Fung HL, Chen S, et al. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science*. 2016; 352:1586–1590. [PubMed: 27339989]
61. Zeng W, Jiang S, Kong X, El-Ali N, Ball AR Jr, Ma CI, Hashimoto N, Yokomori K, Mortazavi A. Single-nucleus RNA-seq of differentiating human myoblasts reveals the extent of fate heterogeneity. *Nucleic Acids Res*. 2016; 44:e158. [PubMed: 27566152]
62. Lv D, Wang X, Dong J, Zhuang Y, Huang S, Ma B, Chen P, Li X, Zhang B, Li Z, et al. Systematic characterization of lncRNAs' cell-to-cell expression heterogeneity in glioblastoma cells. *Oncotarget*. 2016; 7:18403–18414. [PubMed: 26918340]
63. Yunusov D, Anderson L, DaSilva LF, Wysocka J, Ezashi T, Roberts RM, Verjovski-Almeida S. HIPSTR and thousands of lncRNAs are heterogeneously expressed in human embryos, primordial germ cells and stable cell lines. *Sci Rep*. 2016; 6:32753. [PubMed: 27605307]
64. Gambardella G, Carissimo A, Chen A, Cutillo L, Nowakowski TJ, di Bernardo D, Bluelloch R. The impact of microRNAs on transcriptional heterogeneity and gene co-expression across single embryonic stem cells. *Nat Commun*. 2017; 8:14126. [PubMed: 28102192]
65. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, Liu P, Lian Y, Zheng X, Yan J, et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol*. 2013; 20:1131–1139. [PubMed: 23934149]
66. Johnson MB, Wang PP, Atabay KD, Murphy EA, Doan RN, Hecht JL, Walsh CA. Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat Neurosci*. 2015; 18:637–646. [PubMed: 25734491]
67. Kim DH, Marinov GK, Pepke S, Singer ZS, He P, Williams B, Schroth GP, Elowitz MB, Wold BJ. Single-cell transcriptome analysis reveals dynamic changes in lncRNA expression during reprogramming. *Cell Stem Cell*. 2015; 16:88–101. [PubMed: 25575081]
68. Petropoulos S, Edsgard D, Reinius B, Deng Q, Panula SP, Codeluppi S, Reyes AP, Linnarsson S, Sandberg R, Lanner F. Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell*. 2016; 167:285. [PubMed: 27662094]
69. Hu W, Wang T, Yang Y, Zheng S. Tumor heterogeneity uncovered by dynamic expression of long noncoding RNA at single-cell resolution. *Cancer Genet*. 2015; 208:581–586. [PubMed: 26556691]
70. Wang J, Roy B. Single-cell RNA-seq reveals lincRNA expression differences in HeLa-S3 cells. *Biotechnol Lett*. 2017; 39:359–366. [PubMed: 27878653]
71. Lv J, Liu H, Yu S, Liu H, Cui W, Gao Y, Zheng T, Qin G, Guo J, Zeng T, et al. Identification of 4438 novel lincRNAs involved in mouse pre-implantation embryonic development. *Mol Genet Genomics*. 2015; 290:685–697. [PubMed: 25428585]
72. Cabili MN, Dunagin MC, McClanahan PD, Biaisch A, Padovan-Merhar O, Regev A, Rinn JL, Raj A. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome biology*. 2015; 16. [PubMed: 25622821]
73. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublotte JT, Raychowdhury R, Schwartz S, Yosef N, Malboeuf C, Lu D, et al. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature*. 2013; 498:236–240. [PubMed: 23685454]
74. Wu YL, Stubbington MJ, Daly M, Teichmann SA, Rada C. Intrinsic transcriptional heterogeneity in B cells controls early class switching to IgE. *J Exp Med*. 2017; 214:183–196. [PubMed: 27994069]
75. Chen LL. The biogenesis and emerging roles of circular RNAs. *Nat Rev Mol Cell Biol*. 2016; 17:205–211. [PubMed: 26908011]
76. Schmitz SU, Grote P, Herrmann BG. Mechanisms of long noncoding RNA function in development and disease. *Cell Mol Life Sci*. 2016; 73:2491–2509. [PubMed: 27007508]

Figure.