

## Review

## Single Cell Genomics: Advances and Future Perspectives

Iain C. Macaulay<sup>1</sup>, Thierry Voet<sup>1,2\*</sup><sup>1</sup> Single Cell Genomics Centre, Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom, <sup>2</sup> Laboratory of Reproductive Genomics, Department of Human Genetics, KU Leuven, Belgium

**Abstract:** Advances in whole-genome and whole-transcriptome amplification have permitted the sequencing of the minute amounts of DNA and RNA present in a single cell, offering a window into the extent and nature of genomic and transcriptomic heterogeneity which occurs in both normal development and disease. Single-cell approaches stand poised to revolutionise our capacity to understand the scale of genomic, epigenomic, and transcriptomic diversity that occurs during the lifetime of an individual organism. Here, we review the major technological and biological breakthroughs achieved, describe the remaining challenges to overcome, and provide a glimpse into the promise of recent and future developments.

## Introduction

The cell is a fundamental unit of biology, in which the blueprint of the genome is transcribed and translated into biological form and function. Almost all of our current understanding of the genome and its regulation has been derived from studies carried out at the population level—typically thousands or millions of cells analysed in bulk. The resulting analysis, although unquestionably informative, often neglects any heterogeneity that occurs *within* the population of cells.

The genome, despite being widely thought of as stable throughout normal development, has a small probability of acquiring genetic mutations with every cell division [1,2]. Over sufficient divisions, genomic heterogeneity within the organism—known as somatic variation—is a certainty. While such variation lies at the root of many disorders [3,4], including cancer [5], recent studies revealed unexpected levels of genomic variation in normal and diseased tissue, suggesting higher rates of genetic lesion than previously expected [6–12]. Still, little is known about the rate and nature of DNA mutation and how this is influenced by genetic background, lifestyle, and many other factors.

The transcriptome is naturally more dynamic than the genome, reflecting the function—or type—of the cell. There is considerable evidence indicating that cell-to-cell variability in gene expression is ubiquitous, even within a phenotypically “homogeneous” population of cells [13]. The extent of transcriptional heterogeneity and the diversity of cell types in tissues remain, however, largely unknown.

The genomic and transcriptomic composition of individual cells is lost in conventional sequencing studies, which analyse DNA and/or RNA extracted from large populations of cells; and de novo genome mutation and transcriptomic variations in cells will be largely concealed in the bulk signal. Clear insights into many biological processes—from normal development to tumour evolution—will thus only be gained from a detailed understanding of genomic, epigenomic, and transcriptional variation at the single-cell level. Furthermore, some cell types are so rare that single-cell approaches become paramount to their identification and characterisation.

Advances in techniques for the isolation of single cells (Figure 1A), whole genome or transcriptome amplification, and genome-wide analysis platforms—primarily next-generation sequencing (NGS) devices—paved the way for high-resolution analysis of the genome or transcriptome from one cell, which reveals previously obscured biological complexity.

## Single-Cell Whole-Genome Amplification: Methods and Limitations

A diploid human cell contains approximately 7 pg genomic DNA; necessitating amplification prior to microarray- or NGS-based analyses to detect various classes of genetic variation (Figure 1B–1F). Current whole-genome amplification (WGA) principles are based on Multiple Displacement Amplification (MDA), Polymerase Chain Reaction (PCR), or a combination of both (Figure 2A–2C).

Unfortunately, no WGA method is flawless, and their various imperfections can considerably affect the interpretation of the microarray or NGS readout [14]. The breadth of genomic coverage, the amplification bias due to local differences in richness for guanine and cytosine bases (%GC-bias), the prevalence of allelic drop outs (ADO), preferential allelic amplifications (PA), chimeric DNA-molecules, and nucleotide copy errors can vary significantly between different WGA approaches, making some methods better suited than others for detecting specific classes of genetic variation [14–17] (Figure 2D). A comparative analysis of all WGA methods, including the investigation of the advantageous effects of reducing the reaction volume to a nanoliter scale [18,19], against a benchmark case is acute.

## Advances in NGS and Bioinformatics Permit High-Resolution Screening of a Single-Cell Genome

Single-cell WGA products have been analysed using a variety of high-throughput platforms, including DNA-microarrays, SNP-arrays, and NGS (Figure 2D). A key difficulty in the interpretation

**Citation:** Macaulay IC, Voet T (2014) Single Cell Genomics: Advances and Future Perspectives. *PLoS Genet* 10(1): e1004126. doi:10.1371/journal.pgen.1004126

**Editor:** Nancy Maizels, University of Washington, United States of America

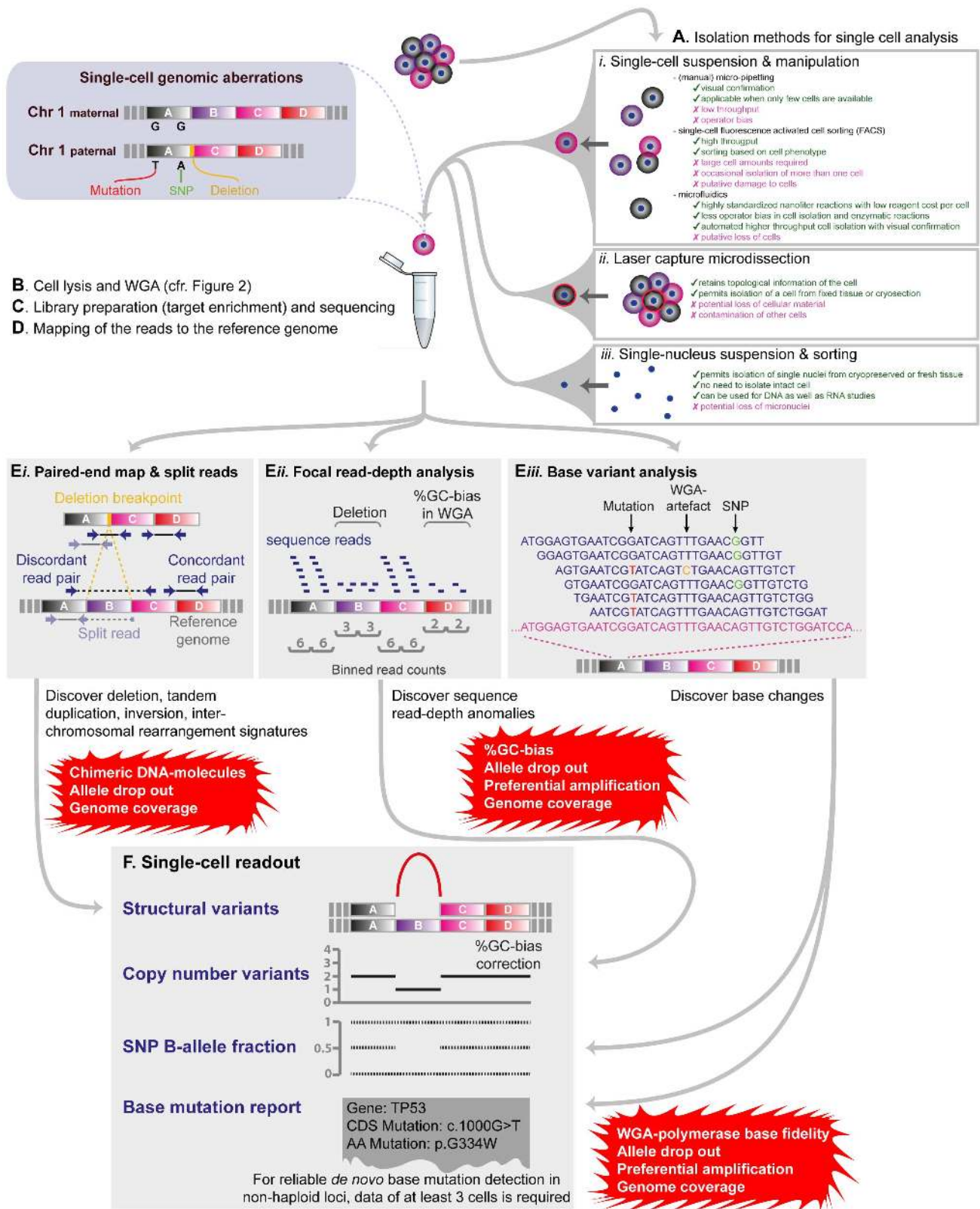
**Published:** January 30, 2014

**Copyright:** © 2014 Macaulay, Voet. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** We acknowledge the Wellcome Trust (UK; <http://www.wellcome.ac.uk/> and <http://www.sanger.ac.uk/>), the Research Foundation Flanders (FWO; <http://www.fwo.be/>; Belgium) [FWO-G.A093.11 to TV], and the KU Leuven [<http://www.kuleuven.be/>; Belgium; SymbioSys, PFV/10/016 to TV]. The funders had no role in the preparation of the article.

**Competing Interests:** I have read the journal's policy and have the following conflicts: TV is a named inventor on submitted patent applications PCT/EP2011/060211; PCT/EP2013/070858; ZL913096.

\* E-mail: [Thierry.Voet@med.kuleuven.be](mailto:Thierry.Voet@med.kuleuven.be)



**Figure 1. Detection of various classes of genetic variation using single-cell WGA-NGS approaches.** A) The most prominent methods for (i–ii) isolating individual cells (including (i) creation of single-cell suspensions—usually by enzymatic tissue disaggregation—and subsequent cell isolation through manual micro-pipetting [37,38,57,105], fluorescence-activated cell sorting [106,107] or microfluidics devices [18,81,108], and (ii) laser capture microdissection [109,110]) as well as (iii) isolating single nuclei [12,32,56,111] are indicated, accompanied with particular advantages and disadvantages. A comprehensive review of single-cell isolation methods is presented by Shapiro et al. [112]. B–D) Subsequently, the cell is lysed and its genome amplified. A standard sequencing library can be prepared from the WGA product for paired-end (or single-end) sequencing. The resulting

(short) sequence reads of the cell are mapped against a reference genome for variant discovery (*Ei–Eiii*). In all steps (*Ei–Eiii* towards *F*), various confounding factors resulting from the WGA process have to be considered in the analysis (indicated in red boxes). *Ei–F* Structural variants can be detected by analysing read-pairs which map discordantly to the reference genome, or by discovering split reads crossing a rearrangement. However, WGA can create various chimeric DNA molecules that resemble structural variants following paired-end sequence analysis of the WGA-product. *Eii–F* Copy number variants are called by “binning” reads that map to particular regions of the genome. By comparing the read count per bin to the counts obtained in a reference sample [17], or an average read count per bin [32], a copy number profile can be calculated. However, single-cell copy number profiles can be distorted by ADO, PA, and %GC-bias during the WGA process. *Eiii–F* Single nucleotide variants (SNVs) can be detected in sequenced single-cell WGA products by aligning the reads with a reference genome. However, three cells carrying the same SNV are required to confidently call the variant.

doi:10.1371/journal.pgen.1004126.g001

of single-cell WGA data on any platform is the separation of the numerous WGA artifacts from the genuine genetic variants present in the template genome.

Standard DNA-microarrays can detect copy number variations (CNVs) larger than 2.5 Mb from a single-cell genome [20–22], while targeted array comparative genomic hybridizations can discover approximately 1 Mb-sized DNA imbalances [23], although remarkably, CNVs as small as 56 kb in single-cell PCR-based WGA products have been detected [24]. Similarly, SNP-arrays can find copy number aberrations encompassing millions of bases in a cell [25–28], but have the advantage of enabling the discovery of copy neutral DNA anomalies and regions of loss-of-heterozygosity (LOH), and allow inferring genome-wide haplotypes [29–31].

NGS has a number of advantages over microarrays enabling improved resolution and accuracy in variant calling [14]. First, NGS can examine every nucleotide amplified from the cell and allows genome-wide discovery of the full spectrum of DNA mutations (Figure 1E), while microarrays only probe for certain CNV loci (Figure 2D). Secondly, sequencing provides digital precision, with one digital unit representing a mapped sequence read. Finally, paired-end sequencing and mapping discloses the linkage between both ends of each linear DNA-molecule in a sequencing library of a single-cell WGA product, allowing the identification of structural variations via read-pairs mapping discordantly to the reference genome (Figure 1E).

Analytical challenges remain in interpreting single-cell NGS data for the full spectrum of genetic variants. Although WGA imperfections due to genome base composition (e.g. %GC-bias) can be computationally corrected for [17,32], the potential for PA and ADO can still generate local distortion in copy number, requiring distinct analyses to distinguish genuine copy number changes from WGA artefacts. Allelic fractions of heterozygous SNPs [26,33,34] or aberrantly mapping read pairs following paired-end sequencing of the WGA product [17] can be used to increase confidence in CNV measurements (Figure 1B–1F). For instance, a real deletion of a diploid locus should show LOH and discordantly mapping read-pairs that explain the DNA loss. Furthermore, the cell cycle stage of the isolated cell must be considered, further complicating the analysis, as cells in S-phase demonstrate a dynamic copy number profile, leading to false structural DNA-imbalance discoveries [35].

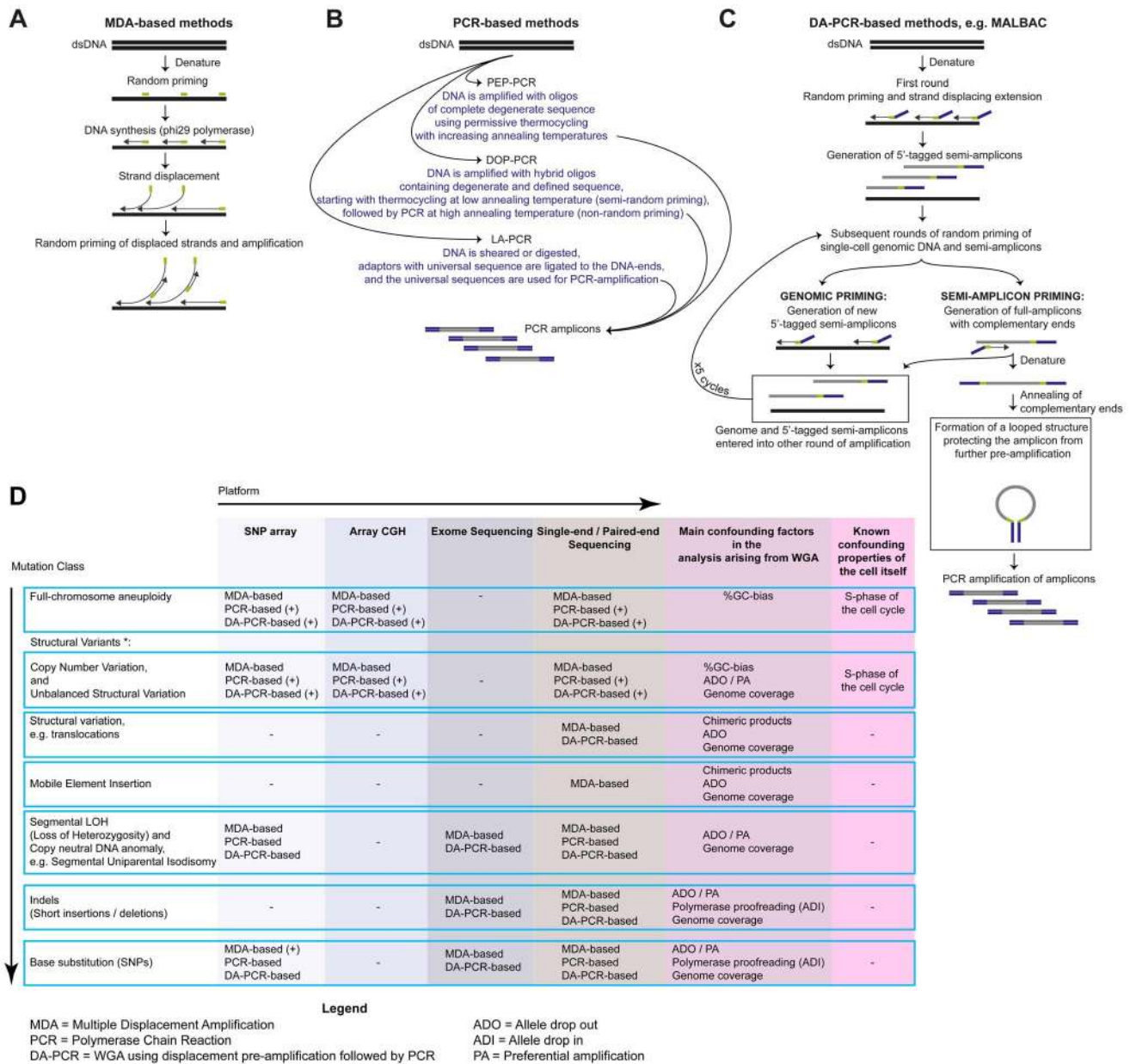
The identification of the full spectrum of intra- and inter-chromosomal (un)balanced structural variants in a single-cell WGA product is still in its infancy—the main difficulty being to filter true structural variants from chimeric DNA generated during WGA, as well as issues with genome coverage (Figure 1Ei, 1F). Although filters have been designed to permit the detection of the structural architecture of DNA copy number variation [17] and even to detect L1-retrotransposition [36], many structural variants are still missed in single-cell analyses. Base alterations, such as SNPs, can be detected in single-cell WGA products (Figure 1Eiii).

However, to call accurate and reliable base substitutions in non-haploid loci, one requires the data of at least three cells to discriminate the variant from a WGA or sequencing error [16,37,38], and as such, detailed characterisation of extremely rare cells or sub-clones within populations may not be possible. Despite these hurdles, several groups have proven the efficacy of single-cell NGS to detect multiple classes of mutation within a genome and even to detect sister chromatid exchanges following single-cell Strand-seq [39]. Step-by-step bioinformatics protocols for analyzing Strand-seq data [40] as well as for copy number profiling single cells through NGS [32] or microarray analysis [34] and commercial solutions (e.g. platforms used within [21,41]) are surfacing.

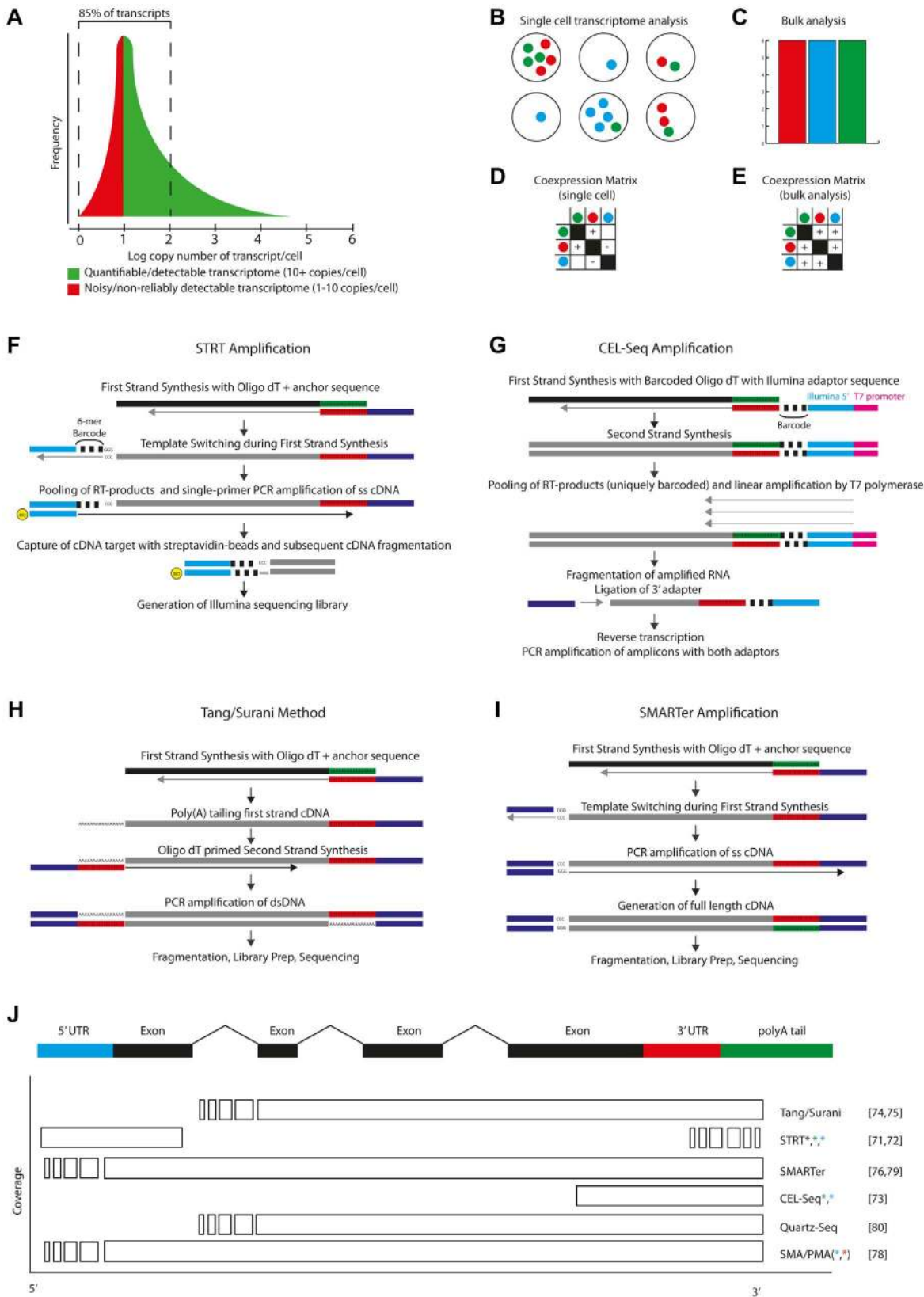
## Single-Cell Genomics Reveals the Extent of Somatic Variation in Development and Disease

The study of multiple classes of mutations at the single-cell level revealed startling insights into the genomic variation that can occur during the human life cycle. Following single-cell genome-wide analysis, up to 7% [18,42] and up to 70% [43–45] of male and female gametes, respectively, contain numerical chromosomal anomalies due to meiotic mis-segregations. Furthermore, sequencing of haploid single sperm cells revealed a base mutation rate of  $2-4 \times 10^{-8}$ —which is severalfold higher than measurements from genome-sequenced pedigree data [46]. Single-cell analyses of human embryos following in vitro fertilization (IVF) demonstrated that the very first cell cycles of human life are prone to numerical and structural chromosome instability [17,25–27,44,47–51]. Various observations indirectly hint that an in vivo conceptus faces a similar period of increased genomic vulnerability [52–55], suggesting that the first cell divisions may represent a spring of DNA mutation, which does not necessarily undermine normal development [8], but can lead to a spectrum of conditions, including loss of conception, genetic disorders, and genetic variation development.

Several studies sequenced and dissected cancer genomes to single-cell resolution, with the aim of understanding tumour development and progression of the disease. Copy number landscapes of single nuclei from primary mammary ductal carcinomas and a paired metastatic liver tumour were generated following low-coverage sequencing. This revealed various chromosomal rearrangements, followed by distinct phases of clonal expansion during tumour evolution and metastasis [56]. Subsequent single-cell exome sequencing studies in bladder [57], kidney [38], and hematopoietic neoplasms [37] provided a detailed characterisation of base mutations in specific genes. Similarly, whole-genome sequencing of multiple MALBAC-amplified cells (Figure 2C) revealed a base mutation rate of a cancer cell line to be increased 10-fold when compared to estimated germ-line ciphers [16]. Furthermore, by sequencing daughter cells of a single mitotic division, the acquisition of new CNVs could be demonstrated for a breast cancer cell line [17].



**Figure 2. Overview of single-cell WGA approaches.** A) Multiple displacement amplification (MDA) initiates with random priming of denatured single-cell DNA template, followed by a 30°C isothermal amplification using a DNA-polymerase with strand-displacement activity, typically phi29 [113]. When the 3'-end of a newly synthesized fragment reaches the 5'-end of an adjoining primed string of nucleotides, it will displace the latter, liberating single-stranded DNA for new primer annealing and DNA-synthesis. B) Primer extension pre-amplification (PEP)-PCR [114], degenerate-oligonucleotide primed (DOP)-PCR [115], and linker-adaptor (LA)-PCR [116] use PCR for WGA. C) WGA methods like PicoPlex [117,118] and MALBAC [16] use displacement pre-amplification to generate PCR-amplifiable fragments (abbreviated as DA-PCR methods here). Specifically, MALBAC initiates with multiple rounds of displacement pre-amplification using a primer that anneals randomly throughout the genome, but contains a specific sequence allowing full amplicons to form looped pre-amplification products of a cell's template DNA. This looping protects previously copied segments from further pre-amplification. Multiple rounds of the displacement pre-amplification reaction, interspersed by a denaturation step, increase the probability that random priming will occur across the genome. D) Classes of genetic variation reported in single cells following WGA and analysis. The proofreading capacity of phi29 improves sequence fidelity during WGA [18,37,38,119]. Furthermore, MDA amplifies the majority of a cell's genome and appears a preferred method for SNP genotyping [15,17,18] or base mutation detection [18,37,38,120], but ADO and PA occurs. Following MDA, single-cell copy number profiles can be distorted [15,17] —although improvements are emerging [19]—and chimeric DNA amplification products are created [17,121]. In general, the (DA-)PCR-based WGA products more accurately preserve the copy number profile of the template genome [15–17] and can be used for SNP genotyping and base mutation detection [16].  
doi:10.1371/journal.pgen.1004126.g002



\* Truncated 3' regions of transcripts have also been detected with STRT

\* Enable multiplexed amplification of single-cell RNA due to early barcoding (see panels F and G)

\* Can retain strand information

\* Can include random primers for first strand cDNA synthesis, whereas all other WTA methods currently rely on oligo-dT priming

**Figure 3. Single-cell transcriptomics.** A) A single cell is thought to contain a few hundred thousand mRNA transcripts, present in a log-normal distribution of abundances, with as much as 85% speculated to be present between 1–100 copies. Current amplification methods are thought to reliably detect transcripts at >5–10 copies per cell. B) Single-cell transcriptomes reveal heterogeneity in gene (co)expression that bulk analysis would not permit. Six single cells are shown, with heterogeneous expression of three genes. C) Bulk analysis would detect uniform expression of all three genes. D–E) Single-cell analysis would reveal underlying heterogeneity but also indicate that two of these genes showed a pattern of co-expression. F) STRT-Seq [71,72] is initiated by reverse transcription using an oligo-dT adaptor-primer. At the 5' cap of the transcript, non-template nucleotides are added by the reverse transcriptase, permitting hybridisation of a barcoded template-switching adaptor-primer. Following pooling of barcoded RT-products, PCR amplification is performed, after which the 5'-ends are captured and sequenced. G) CEL-Seq [73] is initiated using a barcoded oligo-dT primer, which also contains a 5' adaptor sequence and T7 RNA-polymerase priming site. Complementary RNA is generated from the cDNA by T7 RNA polymerase. The cRNA is then fragmented and prepared for (3'-end) paired-end sequencing. H) The Tang/Surani method [74,75], and improved derivatives [69], first generates, then 3' polyadenylates, first strand cDNA. Priming with adaptor-conjugated oligo-dT generates double stranded cDNA, which is then amplified by PCR and sequenced. I) The SMARTer method [76,79] uses template-switching to generate a full-length transcript with adaptor sequences at both ends. These sequences are then used to prime PCR amplification of the transcriptome. The full-length cDNA is used as input for sequencing. J) Overview of the sequence coverage of a typical transcript which would be obtained by each of the currently available single-cell RNA-seq methods [71–80]. doi:10.1371/journal.pgen.1004126.g003

Single-cell genome sequencing continues to provide new insights into genomic (in)stability in various cell types and developmental processes [12]. It will lead to a better understanding of the acquisition of genetic changes during induced pluripotent stem cell derivation and reprogramming, and to insights in the effects of mutagens, carcinogens, ageing, or germline genetic profile on general mutation burden. The methods will enable dissection of the genetic content of individual cells in normal organs, premalignant states, and established tumours, providing insights into the operation of genome maintenance in health and disease.

### Single-Cell Genomics in the Clinic

Single-cell genomics is providing cutting-edge clinical applications, notably in the genetic diagnosis of preimplantation human embryos following IVF (reviewed by Van der Aa et al. [58]). Furthermore, developing embryos shed cells in the maternal blood stream following implantation; the potential to capture and analyse such circulating foetal cells [59,60] may broaden the scope and precision of current non-invasive prenatal testing of circulating foetal DNA in maternal plasma. Single-cell genomics is also applied for studying blood-borne circulating tumour cells (CTCs) [61–63], derived from a solid tumour, to investigate the value of CTCs, in addition to tumour-cell-free DNA [64], for guiding diagnosis, prognosis, and treatment of the cancer.

### Single-Cell Whole-Transcriptome Amplification: Methods and Limitations

A human cell likely contains less than 1 pg of mRNA [65]. Transcripts are thought to be expressed over several orders of magnitude, where many transcripts have low level expression (5–20 transcript copies per cell) [66], with more than 85% having less than 100 copies per cell (Figure 3A). The transcriptome of an individual cell is not fixed, but reflects the functionality of the cell, as well as its responses to acute extrinsic and intrinsic stimuli. In addition to this “regulated” heterogeneity, there is also transcriptional “noise”—heterogeneity which emerges from the kinetics of transcription and mRNA decay between cells within a population. In gene expression analyses of bulk cell populations, it may be impossible to distinguish changes in expression from changes in the cellular composition of the population. Similarly, genes perceived to be co-expressed at the population level may in fact be mutually exclusively expressed when observed at the single-cell level (Figure 3B–3E).

Single-cell reverse transcription and whole-transcriptome amplification (WTA) methods have been developed to permit qPCR [67–69], microarray [70], and more recently, NGS analyses of the

transcriptomes of single cells. Various single-cell RNA-seq methods now exist, each offering an overview of either the 5' end [71,72], 3' end [73] or even the full length [74–80] of transcripts from a single cell (Figure 3F–3J).

Reverse transcription, the initial step in each RNA amplification method, and subsequent conversion of cDNA into amplifiable molecules are likely key limiting factors in the detection and quantification of transcripts in single cells. It is estimated that on average only 5–25% of mRNA-molecules are converted to amplifiable cDNA [72]. Additionally, PCR-based amplification methods have the potential for non-linear amplification, resulting in the distortion of the relative abundance of transcripts. In vitro transcription based WTA-methods, such as CEL-Seq [73] (Figure 3G), may arguably avoid such complications through linear amplification of the transcriptome. Furthermore, nanoliter-scale reactions can demonstrate benefits over microliter-scale processes [81].

### NGS and Bioinformatics Analyses of Single-Cell Transcriptomes

At the most basic level of analysis, a single-cell RNA-seq experiment gives a readout of the abundance of a transcript within a cell. For 3'- or 5'-end sequencing, this is calculated simply as the number of reads mapping to a particular transcript, normalised to the overall number of reads for that cell. If full length RNA is analysed, the number of reads mapping to each transcript is normalised both for the number of reads per cell and, additionally, for transcript length.

Comparative analyses can be applied to measure differences in normalised gene expression between cells. Genes with heterogeneous expression can be identified by their variability within the population; subsequent clustering of variable genes may allow identification of subsets of genes that are co-expressed within a sub-population of cells. Such approaches have been used to dissect specific “bimodal” gene expression patterns within a population of cells [82] and to distinguish co-expression modules in early embryogenesis [83]. While many of the analytical tools for “bulk” mRNA sequencing are also applied for single-cell data, necessary tools specific for single-cell transcriptomics are starting to emerge [72,84,85].

The broad range of transcript abundance in a single cell presents a particular challenge for any amplification method—transcripts present at extremely low levels can still have important biological consequences, and yet, they may be undetectable due to inefficiency of the amplification approach. Even if they are detectable, the influence of technical noise and stochastic effects at these low levels may result in unreliable measurements of relative

abundance within or between individual cells. Thus, a major challenge in quantitative analysis of single-cell transcriptome data is understanding technical noise within or between the samples [86]. The inclusion of RNA spike-ins, such as those developed by the External RNA Controls Consortium (ERCC) [87], can give particular insights into the relative efficiency, detection limits, and technical noise of each amplification method [84,85]. Furthermore, single molecule counting approaches—which incorporate a unique identifier into every molecule prior to amplification—will indicate the extent to which individual RNA molecules are amplified [88].

## Insights from Single-Cell Transcriptomics

Single-cell RNA-seq has already been applied to catalogue allele specific expression and expression of long non-coding RNAs in single blastomeres [74,89] as well as to dissect transcriptional programmes in single cells derived from human and mouse embryos [83,90], revealing insights into the transcriptional modules that are activated at critical points during development. SMARTer WTA [76,79] (Figure 3I) has been used to detect differential exon usage between single cells [82] and to demonstrate a bimodality in gene expression in a phenotypically homogeneous population of bone marrow dendritic cells upon treatment with lipopolysaccharide (LPS) [82]. Here, even genes which were highly expressed were restricted to only a subset of the population—an observation that would have been missed had the population been analysed at a bulk level.

Single-cell transcriptomics has the power to dissect mixed populations of cells; conversely, if only limited material is available, it may permit characterisation of the transcriptome of extremely rare cells, such as CTCs [76].

## The Future: Less Amplification, More Cells, More Types of Data

Many undesirable consequences of WGA and WTA remain to be solved. The ability to reduce [19] or even eliminate amplification of DNA or RNA before sequencing could increase the accuracy and reliability of single-cell analysis. Input requirements for library preparation continue to reduce, and direct library preparation from single-cell genomes has been demonstrated [39,91]. The capacity to directly sequence unamplified DNA and RNA derived from single cells, however, requires further innovation, though direct sequencing of single molecules is already a possibility for DNA and RNA [92,93]. Furthermore, translation of molecular counting principles to single-cell DNA sequence analyses may allow more accurate measurements of CNVs and enable base-error correction [88,94], in addition to haplotyping approaches. Interpreting the full epigenomic status of a cell remains a challenge, but protocols for single-cell DNA-methylation [95–98] and chromosome conformation capture [99] assays are emerging. Excitingly, methodology to analyse both the

(epi)genome and transcriptome of the same cell in parallel is in development and will offer a powerful platform to analyse the exact relationship between genomic variation, regulation, and gene expression.

Typical single-cell sequencing studies have focussed on small numbers of cells (10 s–100 s) but have already demonstrated the potential to distinguish complex heterogeneity at this level. The application of automated cell capture, amplification, and library preparation systems—particularly those utilising nanofluidics approaches—will dramatically increase the scale and affordability of single-cell analyses, such that much larger experiments will emerge.

## Towards a Phylogenetic Tree of a Human Lifetime, and the Discovery of New Cell Types

Studies like the 1000 Genomes Project have contributed greatly to our understanding of genetic and phenotypic variation amongst individuals within a population. However, these studies are grounded on the assumption that the genome of the individual is “fixed” in tissues throughout life.

Considerable evidence is emerging that somatic genomic variation is both common and dynamic in a human being [12,100–104], although little is known about its scale, origin, rate, and nature. Dedicated bioinformatics analyses can extract only the most prevalent heterogeneities (>5% of cells) from populations of cells, representing likely just the tip of the iceberg. To truly understand the full extent of genomic heterogeneity, from conception to death, single-cell genomes must be investigated.

Large-scale single-cell sequencing projects, performed on cells from endodermal, mesodermal, and ectodermal tissues from an individual, will enable construction of a phylogenetic tree of a human lifetime and mapping of the contribution of genetic heterogeneity to the organism. Concurrent single-cell (epi)genomic/transcriptomic studies, on a large enough scale, will allow definitive sub-classification of cell types by gene expression profiles and (epi)genetic status, replacing or enhancing the current schema. Such studies will reveal, in ways that studies of bulk populations cannot address, the relationship between genome sequence, epigenetic status, and gene expression, determining the functional capacity of the cell.

## Conclusion

The last few years have seen rapid development of technologies and methods that permit highly detailed analysis of the genome and transcriptome of a single cell. In parallel, various observations have been made that suggest that both genomic and transcriptomic heterogeneity within an organism may have been considerably underestimated. Single-cell approaches now stand poised to illuminate this new layer of biological complexity during normal development and disease.

## References

1. Nussbaum RL, McInnes RR, Willard HF (2007) *Thompson & Thompson Genetics in Medicine*. Philadelphia: Saunders Elsevier. 600 p.
2. Strachan T, Andrew R (2010) *Human Molecular Genetics*. New York: Garland Science. 781 p.
3. Biesecker LG, Spinner NB (2013) A genomic view of mosaicism and human disease. *Nat Rev Genet* 14: 307–320.
4. Poduri A, Evrony GD, Cai X, Walsh CA (2013) Somatic mutation, genomic variation, and neurological disease. *Science* 341: 1237–1243.
5. Stratton MR, Campbell PJ, Futreal PA (2009) The cancer genome. *Nature* 458: 719–724.
6. De S (2011) Somatic mosaicism in healthy human tissues. *Trends Genet* 27: 217–223.
7. Kloosterman WP, Guryev V, van Roosmalen M, Duran KJ, de Bruijn E, et al. (2011) Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline. *Hum Mol Genet* 20: 1916–1924.
8. Voet T, Vanneste E, Verneesch JR (2011) The human cleavage stage embryo is a cradle of chromosomal rearrangements. *Cytogenet Genome Res* 133: 160–168.
9. Dumanski JP, Piotrowski A (2012) Structural genetic variation in the context of somatic mosaicism. *Methods Mol Biol* 838: 249–272.
10. Kloosterman WP, Cuppen E (2013) Chromothripsis in congenital disorders and cancer: similarities and differences. *Curr Opin Cell Biol* 25: 341–348.
11. Korbel JO, Campbell PJ (2013) Criteria for inference of chromothripsis in cancer genomes. *Cell* 152: 1226–1236.

12. McConnell MJ, Lindberg MR, Brenmand KJ, Piper JC, Voet T, et al. (2013) Mosaic copy number variation in human neurons. *Science* 342: 632–637.
13. Huang S (2009) Non-genetic heterogeneity of cells in development: more than just noise. *Development* 136: 3853–3862.
14. Kumar P, Zamani Esteki M, Van der Aa N, Voet T (2013) How to analyse a single blastomere? Application of whole-genome technologies: micro-arrays and next generation sequencing. In: Sermon K, Viville S, editors. *Textbook of Human Reproductive Genetics*. Cambridge: Cambridge University Press. In press.
15. Treff NR, Su J, Tao X, Northrop LE, Scott RT, Jr. (2011) Single-cell whole-genome amplification technique impacts the accuracy of SNP microarray-based genotyping and copy number analyses. *Mol Hum Reprod* 17: 335–343.
16. Zong C, Lu S, Chapman AR, Xie XS (2012) Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* 338: 1622–1626.
17. Voet T, Kumar P, Van Loo P, Cooke SL, Marshall J, et al. (2013) Single-cell paired-end genome sequencing reveals structural variation per cell cycle. *Nucleic Acids Res* 41: 6119–6138.
18. Wang J, Fan HC, Behr B, Quake SR (2012) Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. *Cell* 150: 402–412.
19. Gole J, Gore A, Richards A, Chiu YJ, Fung HL, et al. (2013) Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. *Nat Biotechnol* 31: 1126–1132.
20. Geigl JB, Obenauf AC, Waldspuehl-Geigl J, Hoffmann EM, Auer M, et al. (2009) Identification of small gains and losses in single cells after whole genome amplification on tiling oligo arrays. *Nucleic Acids Res* 37: e105.
21. Alfarawati S, Fragouli E, Colls P, Wells D (2011) First births after preimplantation genetic diagnosis of structural chromosome abnormalities using comparative genomic hybridization and microarray analysis. *Hum Reprod* 26: 1560–1574.
22. Fiorentino F, Spizzichino L, Bono S, Biricik A, Kokkali G, et al. (2011) PGD for reciprocal and Robertsonian translocations using array comparative genomic hybridization. *Hum Reprod* 26: 1925–1935.
23. Bi W, Berman A, Shaw CA, Stankiewicz P, Gambin T, et al. (2012) Detection of  $>/= 1$  Mb microdeletions and microduplications in a single cell using custom oligonucleotide arrays. *Prenat Diagn* 32: 10–20.
24. Mohlendick B, Bartenhagen C, Behrens B, Honisch E, Raba K, et al. (2013) A Robust Method to Analyze Copy Number Alterations of Less than 100 kb in Single Cells Using Oligonucleotide Array CGH. *PLoS One* 8: e67031. doi:10.1371/journal.pone.0067031
25. Vanneste E, Voet T, Le Caignec C, Ampe M, Konings P, et al. (2009) Chromosome instability is common in human cleavage-stage embryos. *Nat Med* 15: 577–583.
26. Johnson DS, Gemelos G, Baner J, Ryan A, Cinnioğlu C, et al. (2010) Preclinical validation of a microarray method for full molecular karyotyping of blastomeres in a 24-h protocol. *Hum Reprod* 25: 1066–1075.
27. Voet T, Vanneste E, Van der Aa N, Melotte C, Jackmaert S, et al. (2011) Breakage-fusion-bridge cycles leading to inv dup del occur in human cleavage stage embryos. *Hum Mutat* 32: 783–793.
28. van Uum CM, Stevens SJ, Dreesen JC, Drusedau M, Smeets HJ, et al. (2012) SNP array-based copy number and genotype analyses for preimplantation genetic diagnosis of human unbalanced translocations. *Eur J Hum Genet* 20: 938–944.
29. Handyside AH, Harton GL, Mairani B, Thornhill AR, Affara N, et al. (2010) Karyomapping: a universal method for genome wide analysis of genetic disease based on mapping crossovers between parental haplotypes. *J Med Genet* 47: 651–658.
30. Fan HC, Wang J, Potanina A, Quake SR (2011) Whole-genome molecular haplotyping of single cells. *Nat Biotechnol* 29: 51–57.
31. Altarescu G, Zeevi DA, Zeligson S, Perlberg S, Eldar-Geva T, et al. (2013) Familial haplotyping and embryo analysis for Preimplantation Genetic Diagnosis (PGD) using DNA microarrays: a proof of principle study. *J Assist Reprod Genet* 30: 1595–1603.
32. Baslan T, Kendall J, Rodgers L, Cox H, Riggs M, et al. (2012) Genome-wide copy number analysis of single cells. *Nat Protoc* 7: 1024–1041.
33. Treff NR, Su J, Tao X, Levy B, Scott RT, Jr. (2010) Accurate single cell 24 chromosome aneuploidy screening using whole genome amplification and single nucleotide polymorphism microarrays. *Fertil Steril* 94: 2017–2021.
34. Konings P, Vanneste E, Jackmaert S, Ampe M, Verbeke G, et al. (2012) Microarray analysis of copy number variation in single cells. *Nat Protoc* 7: 281–310.
35. Van der Aa N, Cheng J, Mateiu L, Esteki MZ, Kumar P, et al. (2013) Genome-wide copy number profiling of single cells in S-phase reveals DNA-replication domains. *Nucleic Acids Res* 41: e66.
36. Evrony GD, Cai X, Lee E, Hills LB, Elhosary PC, et al. (2012) Single-neuron sequencing analysis of 11 retrotransposition and somatic mutation in the human brain. *Cell* 151: 483–496.
37. Hou Y, Song L, Zhu P, Zhang B, Tao Y, et al. (2012) Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. *Cell* 148: 873–885.
38. Xu X, Hou Y, Yin X, Bao L, Tang A, et al. (2012) Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* 148: 886–895.
39. Falconer E, Hills M, Naumann U, Poon SS, Chavez EA, et al. (2012) DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. *Nat Methods* 9: 1107–1112.
40. Hills M, O'Neill K, Falconer E, Brinkman R, Lansdorp PM (2013) BAIT: Organizing genomes and mapping rearrangements in single cells. *Genome Med* 5: 82.
41. Mertzaniidou A, Spits C, Nguyen HT, Van de Velde H, Sermon K (2013) Evolution of aneuploidy up to Day 4 of human preimplantation development. *Hum Reprod* 28: 1716–1724.
42. Lu S, Zong C, Fan W, Yang M, Li J, et al. (2012) Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. *Science* 338: 1627–1630.
43. Handyside AH, Montag M, Magli MC, Repping S, Harper J, et al. (2012) Multiple meiotic errors caused by predivision of chromatids in women of advanced maternal age undergoing in vitro fertilisation. *Eur J Hum Genet* 20: 742–747.
44. Nagaoka SI, Hassold TJ, Hunt PA (2012) Human aneuploidy: mechanisms and new insights into an age-old problem. *Nat Rev Genet* 13: 493–504.
45. Hou Y, Fan W, Yan L, Li R, Lian Y, et al. (2013) Genome analyses of single human oocytes. *Cell* 155: 1492–1506.
46. Conrad DF, Keebler JE, DePristo MA, Lindsay SJ, Zhang Y, et al. (2011) Variation in genome-wide mutation rates within and between human families. *Nat Genet* 43: 712–714.
47. Vanneste E, Voet T, Melotte C, Debrock S, Sermon K, et al. (2009) What next for preimplantation genetic screening? High mitotic chromosome instability rate provides the biological basis for the low success rate. *Hum Reprod* 24: 2679–2682.
48. Treff NR, Levy B, Su J, Northrop LE, Tao X, et al. (2010) SNP microarray-based 24 chromosome aneuploidy screening is significantly more consistent than FISH. *Mol Hum Reprod* 16: 583–589.
49. Gutierrez-Mateo C, Colls P, Sanchez-Garcia J, Escudero T, Prates R, et al. (2011) Validation of microarray comparative genomic hybridization for comprehensive chromosome analysis of embryos. *Fertil Steril* 95: 953–958.
50. Treff NR, Su J, Taylor D, Scott RT, Jr. (2011) Telomere DNA deficiency is associated with development of human embryonic aneuploidy. *PLoS Genet* 7: e1002161. doi:10.1371/journal.pgen.1002161
51. Alfarawati S, Fragouli E, Colls P, Wells D (2012) Embryos of robertsonian translocation carriers exhibit a mitotic interchromosomal effect that enhances genetic instability during early development. *PLoS Genet* 8: e1003025. doi:10.1371/journal.pgen.1003025
52. Pflueger S (1999) *Cytogenetics of Spontaneous Abortion*. In: Gersen S, Keagle M, editors. *The Principles of Clinical Cytogenetics*. New Jersey: Humana Press. pp. 317–343.
53. Macklon NS, Geraedts JP, Fauser BC (2002) Conception to ongoing pregnancy: the 'black box' of early pregnancy loss. *Hum Reprod Update* 8: 333–343.
54. Santos MA, Kuijk EW, Macklon NS (2010) The impact of ovarian stimulation for IVF on the developing embryo. *Reproduction* 139: 23–34.
55. Robberecht C, Voet T, Zamani Esteki M, Nowakowska BA, Vermeesch JR (2013) Nonallelic homologous recombination between retrotransposable elements is a driver of de novo unbalanced translocations. *Genome Res* 23: 411–418.
56. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, et al. (2011) Tumour evolution inferred by single-cell sequencing. *Nature* 472: 90–94.
57. Li Y, Xu X, Song L, Hou Y, Li Z, et al. (2012) Single-cell sequencing analysis characterizes common and cell-lineage-specific mutations in a muscle-invasive bladder cancer. *Gigascience* 1: 12.
58. Van der Aa N, Zamani Esteki M, Vermeesch JR, Voet T (2013) Preimplantation genetic diagnosis guided by single-cell genomics. *Genome Med* 5: 71.
59. Elias S, Price J, Dockter M, Wachtel S, Tharapel A, et al. (1992) First trimester prenatal diagnosis of trisomy 21 in fetal cells from maternal blood. *Lancet* 340: 1033.
60. Simpson JL (2013) Cell-free fetal DNA and maternal serum analytes for monitoring embryonic and fetal status. *Fertil Steril* 99: 1124–1134.
61. Heitzer E, Auer M, Gasch C, Pichler M, Ulz P, et al. (2013) Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. *Cancer Res* 73: 2965–2975.
62. Ni X, Zhuo M, Su Z, Duan J, Gao Y, et al. (2013) Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. *Proc Natl Acad Sci U S A* 110: 21083–21088.
63. Swenhuis JF, Reumers J, Thys K, Aerssens J, Terstappen LW (2013) Efficiency of whole genome amplification of Single Circulating Tumor Cells enriched by CellSearch and sorted by FACS. *Genome Med* 5: 106.
64. Navin N, Hicks J (2011) Future medical applications of single-cell sequencing in cancer. *Genome Med* 3: 31.
65. Kawasaki ES (2004) Microarrays and the gene expression profile of a single cell. *Ann N Y Acad Sci* 1020: 92–100.
66. Subkhankulova T, Gilchrist MJ, Livesey FJ (2008) Modelling and measuring single cell RNA expression levels find considerable transcriptional differences among phenotypically identical cells. *BMC Genomics* 9: 268.
67. Warren L, Bryder D, Weissman IL, Quake SR (2006) Transcription factor profiling in individual hematopoietic progenitors by digital RT-PCR. *Proc Natl Acad Sci U S A* 103: 17807–17812.



68. Guo G, Huss M, Tong GQ, Wang C, Li Sun L, et al. (2010) Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev Cell* 18: 675–685.
69. Huang H, Goto M, Tsunoda H, Sun L, Taniguchi K, et al. (2013) Non-biased and efficient global amplification of a single-cell cDNA library. *Nucleic Acids Res*. In press.
70. Iscove NN, Barbara M, Gu M, Gibson M, Modi C, et al. (2002) Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA. *Nat Biotechnol* 20: 940–943.
71. Islam S, Kjallquist U, Moliner A, Zajac P, Fan JB, et al. (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res* 21: 1160–1167.
72. Islam S, Kjallquist U, Moliner A, Zajac P, Fan JB, et al. (2012) Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nat Protoc* 7: 813–828.
73. Hashimshony T, Wagner F, Sher N, Yanai I (2012) CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep* 2: 666–673.
74. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, et al. (2009) mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 6: 377–382.
75. Tang F, Barbacioru C, Nordman E, Li B, Xu N, et al. (2010) RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat Protoc* 5: 516–535.
76. Ramskold D, Luo S, Wang YC, Li R, Deng Q, et al. (2012) Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol* 30: 777–782.
77. Kohn AB, Moroz TP, Barnes JP, Netherton M, Moroz LL (2013) Single-cell semiconductor sequencing. *Methods Mol Biol* 1048: 247–284.
78. Pan X, Durrett RE, Zhu H, Tanaka Y, Li Y, et al. (2013) Two methods for full-length RNA sequencing for low quantities of cells and single cells. *Proc Natl Acad Sci U S A* 110: 594–599.
79. Picelli S, Bjorklund AK, Faridani OR, Sagasser S, Winberg G, et al. (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods* 10: 1096–1098.
80. Sasagawa Y, Nikaido I, Hayashi T, Danno H, Uno KD, et al. (2013) Quartz-Seq: a highly reproducible and sensitive single-cell RNA-Seq reveals non-genetic gene expression heterogeneity. *Genome Biol* 14: R31.
81. Wu AR, Neff NF, Kalisky T, Dalerba P, Treutlein B, et al. (2013) Quantitative assessment of single-cell RNA-sequencing methods. *Nat Methods* 11: 41–46.
82. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaubomme JT, et al. (2013) Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* 498: 236–240.
83. Xue Z, Huang K, Cai C, Cai L, Jiang CY, et al. (2013) Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* 500: 593–597.
84. Brennecke P, Anders S, Kim JK, Kolodziejczyk AA, Zhang X, et al. (2013) Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods* 10: 1093–1095.
85. Katayama S, Tohonen V, Linnarsson S, Kere J (2013) SAMstr: statistical test for differential expression in single-cell transcriptome with spike-in normalization. *Bioinformatics* 29: 2943–2945.
86. Marinov GK, Williams BA, McCue K, Schroth GP, Gertz J, et al. (2013) From single-cell to cell-pool transcriptomes: stochasticity in gene expression and RNA splicing. *Genome Res*. In press.
87. Jiang L, Schlesinger F, Davis CA, Zhang Y, Li R, et al. (2011) Synthetic spike-in standards for RNA-seq experiments. *Genome Res* 21: 1543–1551.
88. Kivioja T, Vaharautio A, Karlsson K, Bonke M, Enge M, et al. (2012) Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods* 9: 72–74.
89. Tang F, Barbacioru C, Nordman E, Bao S, Lee C, et al. (2011) Deterministic and stochastic allele specific gene expression in single mouse blastomeres. *PLoS One* 6: e21208. doi:10.1371/journal.pone.0021208
90. Yan L, Yang M, Guo H, Yang L, Wu J, et al. (2013) Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol* 20: 1131–1139.
91. Falconer E, Lansdorp PM (2013) Strand-seq: A unifying tool for studies of chromosome segregation. *Semin Cell Dev Biol* 24: 643–652.
92. Ozsolak F, Platt AR, Jones DR, Reifemberger JG, Sass LE, et al. (2009) Direct RNA sequencing. *Nature* 461: 814–818.
93. Coupland P, Chandra T, Quail M, Reik W, Swerdlow H (2012) Direct sequencing of small genomes on the Pacific Biosciences RS without library preparation. *Biotechniques* 53: 365–372.
94. Schmitt MW, Kennedy SR, Salk JJ, Fox EJ, Hiatt JB, et al. (2012) Detection of ultra-rare mutations by next-generation sequencing. *Proc Natl Acad Sci U S A* 109: 14508–14513.
95. El Hajj N, Trapphoff T, Linke M, May A, Hansmann T, et al. (2011) Limiting dilution bisulfite (pyro)sequencing reveals parent-specific methylation patterns in single early mouse embryos and bovine oocytes. *Epigenetics* 6: 1176–1188.
96. Denomme MM, Zhang L, Mann MR (2012) Single oocyte bisulfite mutagenesis. *J Vis Exp*. doi: 10.3791/4046.
97. Guo H, Zhu P, Wu X, Li X, Wen L, et al. (2013) Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res* 23: 2126–2135.
98. Lorthongpanich C, Cheow LF, Balu S, Quake SR, Knowles BB, et al. (2013) Single-cell DNA-methylation analysis reveals epigenetic chimerism in preimplantation embryos. *Science* 341: 1110–1112.
99. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, et al. (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502: 59–64.
100. Abyzov A, Mariani J, Palejev D, Zhang Y, Hancu MS, et al. (2012) Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature* 492: 438–442.
101. Jacobs KB, Yeager M, Zhou W, Wacholder S, Wang Z, et al. (2012) Detectable clonal mosaicism and its relationship to aging and cancer. *Nat Genet* 44: 651–658.
102. Laurie CC, Laurie CA, Rice K, Doherty KF, Zelnick LR, et al. (2012) Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet* 44: 642–650.
103. O'Huallachain M, Karczewski KJ, Weissman SM, Urban AE, Snyder MP (2012) Extensive genetic variation in somatic human tissues. *Proc Natl Acad Sci U S A* 109: 18018–18023.
104. Bonnefond A, Skrobek B, Lobbens S, Eury E, Thuillier D, et al. (2013) Association between large detectable clonal mosaicism and type 2 diabetes with vascular complications. *Nat Genet* 45: 1040–1043.
105. Kirkness EF, Grindberg RV, Yee-Greenbaum J, Marshall CR, Scherer SW, et al. (2013) Sequencing of isolated sperm cells for direct haplotyping of a human genome. *Genome Res* 23: 826–832.
106. Dalerba P, Kalisky T, Sahoo D, Rajendran PS, Rothenberg ME, et al. (2011) Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat Biotechnol* 29: 1120–1127.
107. Potter NE, Ermini L, Papaemmanuil E, Cazzaniga G, Vijayaraghavan G, et al. (2013) Single-cell mutational profiling and clonal phylogeny in cancer. *Genome Res* 23: 2115–2125.
108. Lecault V, White AK, Singhal A, Hansen CL (2012) Microfluidic single cell analysis: from promise to practice. *Curr Opin Chem Biol* 16: 381–390.
109. Frumkin D, Wasserstrom A, Itzkovitz S, Harmelin A, Rechavi G, et al. (2008) Amplification of multiple genomic loci from single cells isolated by laser microdissection of tissues. *BMC biotechnology* 8: 17.
110. Boone DR, Sell SL, Hellmich HL (2013) Laser capture microdissection of enriched populations of neurons or single neurons for gene expression analysis after traumatic brain injury. *J Vis Exp*. doi: 10.3791/50308.
111. Grindberg RV, Yee-Greenbaum JL, McConnell MJ, Novotny M, O'Shaughnessy AL, et al. (2013) RNA-sequencing from single nuclei. *Proc Natl Acad Sci U S A* 110: 19802–19807.
112. Shapiro E, Biezuner T, Linnarsson S (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet* 14: 618–630.
113. Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, et al. (2002) Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci U S A* 99: 5261–5266.
114. Zhang L, Cui X, Schmitt K, Hubert R, Navidi W, et al. (1992) Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci U S A* 89: 5847–5851.
115. Cheung VG, Nelson SF (1996) Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. *Proc Natl Acad Sci U S A* 93: 14676–14679.
116. Klein CA, Schmidt-Kittler O, Schardt JA, Pantel K, Speicher MR, et al. (1999) Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. *Proc Natl Acad Sci U S A* 96: 4494–4499.
117. Langmore JP (2002) Rubicon Genomics, Inc. *Pharmacogenomics* 3: 557–560.
118. Blainey PC (2013) The future is now: single-cell genomics of bacteria and archaea. *FEMS Microbiol Rev* 37: 407–427.
119. Esteban JA, Salas M, Blanco L (1993) Fidelity of phi 29 DNA polymerase. Comparison between protein-primed initiation and DNA polymerization. *J Biol Chem* 268: 2719–2726.
120. Gundry M, Li W, Maqbool SB, Vijg J (2012) Direct, genome-wide assessment of DNA mutations in single cells. *Nucleic Acids Res* 40: 2032–2040.
121. Lasken RS, Stockwell TB (2007) Mechanism of chimera formation during the Multiple Displacement Amplification reaction. *BMC Biotechnol* 7: 19.