levels in muscle cells to permit adaptive changes in gene expression in response to environmental stimuli.

The exciting findings of Bodega *et al.* have revealed an important new layer of regulation for PRC2-mediated transcriptional repression that allows postmitotic cells to respond to cellular stress. While the Ezh1β transcript was detected in a large number of adult tissues[12], it remains to be seen whether Eed sequestration modulates transcriptional repression in other cell types.

ACKNOWLEDGMENTS
Work in the Brand and Dilworth labs is supported by grants from the Canadian Institutes of Health Research (MOP-82813 to M.B. and FDN-143330 to F.J.D.).

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Geisler, S.J. & Paro, R. *Development* **142**, 2876–2887 (2015).
2. Pasini, D. & Di Croce, L. *Curr. Opin. Genet. Dev.* **36**, 50–58 (2016).
3. Shen, X. *et al. Mol. Cell* **32**, 491–502 (2008).
4. Cao, R. & Zhang, Y. *Mol. Cell* **15**, 57–67 (2004).
5. Margueron, R. *et al. Nature* **461**, 762–767 (2009).
6. Ezhkova, E. *et al. Genes Dev.* **25**, 485–498 (2011).
7. O'Carroll, D. *et al. Mol. Cell. Biol.* **21**, 4330–4336 (2001).
8. Margueron, R. *et al. Mol. Cell* **32**, 503–518 (2008).
9. Mu, W., Starmer, J., Shibata, Y., Yee, D. & Magnuson, T. *Dev. Biol.* **424**, 198–207 (2017).
10. Caretti, G., Di Padova, M., Micales, B., Lyons, G.E. & Sartorelli, V. *Genes Dev.* **18**, 2627–2638 (2004).
11. Stojic, L. *et al. Epigenetics Chromatin* **4**, 16 (2011).
12. Bodega, B. *et al. Nat. Struct. Mol. Biol.* **24**, 444–452 (2017).
13. Xu, B. *et al. Blood* **125**, 346–357 (2015).
14. Mousavi, K., Zare, H., Wang, A.H. & Sartorelli, V. *Mol. Cell* **45**, 255–262 (2012).

Capturing heterogeneity: single-cell structures of the 3D genome

Elzo de Wit

One of the striking features of cells seen through a microscope is the heterogeneous organization of the nuclei. A combination of molecular methods and computational modeling has now been used to reconstruct accurate 3D structures of the genome inside single nuclei.

Single-cell genomics methods are quickly revolutionizing our understanding of cell-to-cell variability. Single-cell RNA sequencing (scRNA-seq) has led to the probing of tissue heterogeneity at the gene expression level with unprecedented resolution[1]. Sequencing DNA from a single cell has enabled lineage reconstruction and revealed genetic variability in tumor samples, characterizing tumorigenesis at the DNA level[2]. Furthermore, scATAC-seq and scDamID make it possible to chart the chromatin landscape at the single-cell level[3–5]. Together, these and other methods are illuminating the substantial heterogeneity present in cellular populations and enable the reconstruction of tissues and the probing of interactions between cells within these tissues.

All nuclei within heterogeneous samples carry the same genome, and therefore differences between cells must be the result of epigenetic differences. One can view the organization of the genome inside the nucleus (or the 3D genome) as an additional layer of epigenetic regulation. 3D genomes can be subdivided into active (A) and inactive (B) compartments[6] that are strongly correlated with their radial position in the nucleus. Inactive regions are located at the nuclear periphery, overlapping with nuclear lamina-associated domains (LADs)[7]. Active regions, on the other hand, are often found more internally. A and B compartments can be further subdivided into topologically associating domains (TADs)[8,9], which are insulated genetic neighborhoods that ensure proper communication between enhancers and promoters within the same TAD and prevent miscommunication of enhancers and promoters that lie in different TADs[10]. Within TADs, an additional layer of complexity is created by preferential looping of sites bound by the insulator protein CTCF[11].

The above discoveries have been made possible by the development of Hi-C[7], a molecular method that enables the identification of genomic regions that are in close proximity. Whereas most Hi-C analyses are done on cell populations, typically using millions of cells, in 2013, Nagano *et al.* described a method for downscaling Hi-C to single cells[12]. The recent work by Stevens *et al.*[13] introduces two critical steps that make this analysis amenable to high-resolution computational modeling of the spatial organization of a complete genome. First, the authors employ live-cell microscopy to select cells in the G1 phase of the cell cycle. In addition, they use haploid embryonic stem cells. This is crucial, because in a diploid genome, mapped interactions derive from both parental chromosomes, so their complement is a mix of two differently folded chromosomes. In a haploid genome, such ambiguities do not exist.

The authors have also developed a computational analysis framework to perform 3D modeling of entire genomes. The interactions that are found in a single nucleus by Hi-C are used as restraints to generate 3D models of chromosomes (**Fig. 1a**). In the Hi-C data from haploid cells, very few interactions are detected that would not be possible in the 3D genome models (for example, between loci that are far apart in the model), strongly suggesting the models are correct. When a 3D genome model was created using data from a single cell containing a duplicated chromosome, the resulting 3D structure of this chromosome contained many interactions that violated the model, showing that the model for this chromosome was incorrect. This underscores the importance of using haploid genomes for single-cell Hi-C when the objective is to model chromosomes.

To further validate the observed structures, the authors project well-known measures of genome organization onto the 3D models, such as A/B compartments, LADs and gene expression. The resolved models clearly recapitulate previous observations regarding genome organization: the LADs and the B compartment are found at the periphery of the chromosome territory, whereas active genomic regions are buried inside the chromosomes and therefore the nucleus (**Fig. 1b**). Deep within the nucleus, a second layer of B compartment and LADs are found lining the nucleolus, which is in agreement with previous single-cell tracking experiments[14]. At a deeper level, the models also provide a measure for the compaction of TADs. In different cells, TADs assume different levels of compaction, suggesting that TADs are not inert globules with very high self-interaction. Rather, they are dynamic structures that can assume distinct conformations (**Fig. 1c**). These observations are in line with the loop-extrusion model[15], which has recently been proposed based on computational simulations as an explanation for TAD formation[16] (**Fig. 1d**).

Division of Gene Regulation, Netherlands Cancer Institute, Amsterdam, the Netherlands.
e-mail: e.d.wit@nki.nl

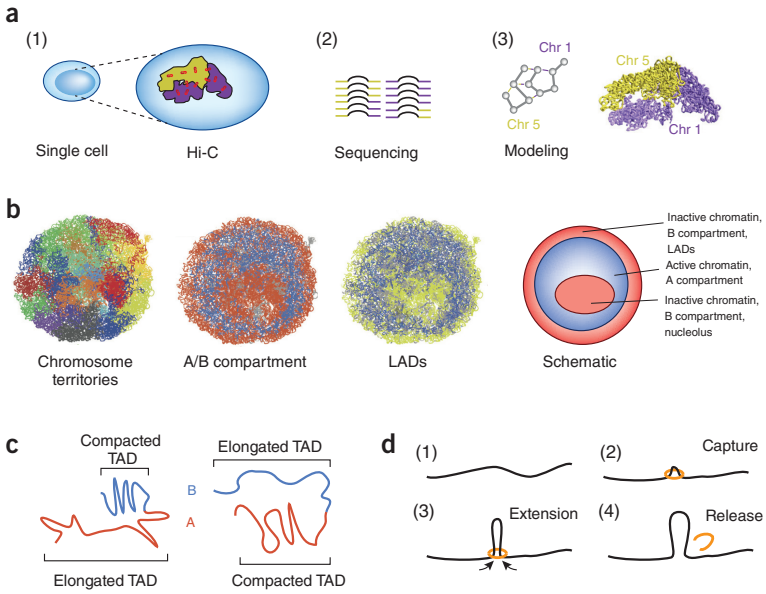

Figure 1 Single-cell models of chromosomes reveal principle characteristics of nuclear organization. (**a**) Single-cell Hi-C is performed to identify contacts between chromosomal regions (1). The contacts are sequenced using Illumina paired-end sequencing (2) and used as restraints in performing computational modeling of 3D genome structures (3). (**b**) Genome structures recapitulate known features of nuclear organization. Images of genome structures are adapted from ref. 13. (**c**) Structures of TADs show that they can exist in both compacted and elongated conformations. (**d**) Schematic representation of the loop-extrusion model. A chromatin region is captured by an extrusion complex forming a tiny loop (2), which is actively extended, leading to the formation of larger loops (3). Ultimately, the extrusion complex releases the chromatin, leading to the dissolving of the loop (4).

Stevens *et al.*[13] are able to resolve 3D models using a relatively modest number of unique contacts (at least 26,000), which gives a resolution of ~100 kb. To achieve the resolution of chromatin interactions mediated by CTCF, the resolution of the model needs to be at least an order of magnitude higher, which requires improvement of the recovery of contacts from Hi-C data. This will be a major challenge for future experiments.

An important conclusion is that diploid- (or population-) based Hi-C data cannot be used to resolve 3D genome models. This also impacts existing 3D genome models[17] and modeling software. It will be interesting to see how population-based chromosome models stack up to single-cell models and how single-cell data can be used to improve the population-based modeling. The requirement of haploid cells severely limits the choice of model systems for this type of modeling experiment. An alternative would be to use model systems with extremely high genetic diversity between the parental chromosomes, such as *Mus musculus* × *Mus castaneus* F_1 hybrid embryonic stem cells, which on average have a sequence variant every 130 bp. This would enable unambiguous assignment of contacts to one of the two parental chromosomes (**Fig. 2a**). For obvious reasons, this is not a solution for human cells, in which the genetic diversity between alleles is ~0.1% or 1/1,000 bp. However, long-read technology[18] coupled to fully haplotyped genomes may offer a solution here (**Fig. 2b**).

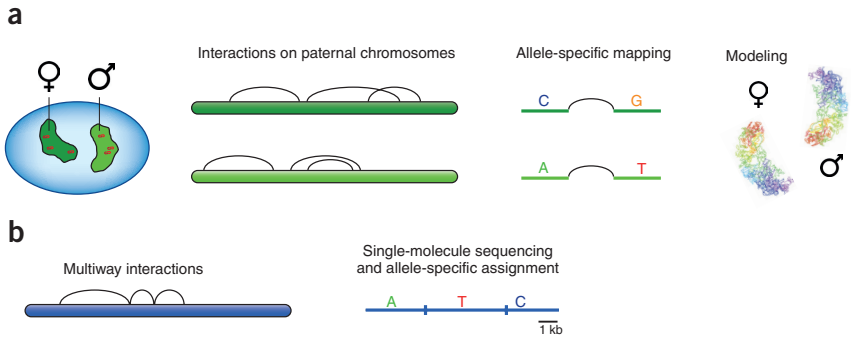

Figure 2 Alternatives to using haploid genomes for single-cell genome modeling. (**a**) In a diploid genome, two copies of every chromosome are present, each with its own specific conformation and contacts. If homologous chromosomes have a high enough sequence diversity, read pairs, and therefore contacts, can be assigned to the maternal or paternal chromosomes. This information can then be used to perform modeling on the complete diploid chromosome set. Images of genome structures adapted from ref. 13. (**b**) For genomes with lower sequence diversity, larger fragments containing multiple contacts should be sequenced. When the fragment contains a genetic variant, this fragment can be assigned to a parental chromosome. Given that the human genome contains approximately one sequence variant every 1 kb, sequencing a 10-kb fragment should yield roughly ten informative variants.

Stevens *et al.* analyzed eight single cells[13]. However, the number of possible constellations of nuclear interactions also shown in this work makes the organization of chromosomes very heterogeneous by definition. Therefore, to draw meaningful quantitative conclusions about cell-to-cell heterogeneity, it is necessary to dramatically increase the throughput. Recently, an approach using combinatorial barcoding has analyzed thousands of cells in parallel[19], but this method resulted in a lower read depth compared to Stevens *et al.*[13] The authors of the original single-cell Hi-C analysis have shown that robotics can also be used to analyze thousands of cells[20]. When single-cell genome modeling is performed with strongly increased throughput, it will be possible to start systematically cataloguing the heterogeneity in chromosome models and see whether common organizational themes can be identified. It will be particularly interesting to see if the cell-to-cell heterogeneity that is identified in scRNA-seq is reflected in 3D genome models or whether the 3D genome is a scaffold upon which cell-to-cell variability plays out.

COMPETING FINANCIAL INTERESTS
The author declares no competing financial interests.

1. Kolodziejczyk, A.A., Kim, J.K., Svensson, V., Marioni, J.C. & Teichmann, S.A. *Mol. Cell* **58**, 610–620 (2015).
2. Shapiro, E., Biezuner, T. & Linnarsson, S. *Nat. Rev. Genet.* **14**, 618–630 (2013).
3. Buenrostro, J.D. *et al. Nature* **523**, 486–490 (2015).
4. Cusanovich, D. A. *et al. Science* **348**, 910–914 (2015).
5. Kind, J. *et al. Cell* **163**, 134–147 (2015).
6. Lieberman-Aiden, E. *et al. Science* **326**, 289–293 (2009).
7. Peric-Hupkes, D. *et al. Mol. Cell* **38**, 603–613 (2010).
8. Dixon, J.R. *et al. Nature* **485**, 376–380 (2012).
9. Nora, E.P. *et al. Nature* **485**, 381–385 (2012).
10. Flavahan, W.A. *et al. Nature* **529**, 110–114 (2016).
11. Rao, S.S.P. *et al. Cell* **159**, 1665–1680 (2014) .
12. Nagano, T. *et al. Nature* **502**, 59–64 (2013).
13. Stevens, T.J. *et al. Nature* **544**, 59–64 (2017).
14. Kind, J. *et al. Cell* **153**, 178–192 (2013).
15. Nasmyth, K. *Annu. Rev. Genet.* **35**, 673–745 (2001).
16. Fudenberg, G. *et al. Cell Rep.* **15**, 2038–2049 (2016).
17. Varoquaux, N., Ay, F., Noble, W.S. & Vert, J.-P. *Bioinformatics* **30**, i26–i33 (2014).
18. Metzker, M.L. *Nat. Rev. Genet.* **11**, 31–46 (2010).
19. Ramani, V. *et al. Nat. Methods* **14**, 263–266 (2017).
20. Nagano, T. *et al.* Preprint at https://doi.org/10.1101/094466 (2016).