with the interpretation of haplotype information from low-depth sequencing data (i.e., haplotype blocks based on shared ancestry[10]), which enables correct phasing of ~80% of the cell line's heterozygous SNPs.

Although these new approaches[1–3] show great promise, several limitations should be considered. First, their performance depends not only on the depth of coverage achieved for the Hi-C experiments but also on the lengths of the DNA reads themselves, because shorter reads are more likely to map to multiple locations in the genome and are therefore unlikely to be interpretable. Regions of the genome with high-repeat content, which include segmental duplications, interspersed repeats and even known protein-coding exons, remain difficult to interpret with short-read sequencing technologies.

Second, although proximity ligation shows great promise for identifying long-range DNA interactions, Hi-C signals are better suited to derive information about the order, rather than orientation, of segments along the chromosome. Notably, Burton *et al.*[1] report an enrichment of segmental duplications and simple repeats among the scaffolds that were ordered incorrectly by LACHESIS. A significant portion of these errors involve incorrect contig orientation rather than incorrect contig order. By comparison, the ability to phase SNPs with HaploSeq depends on the genomic density of heterozygous genetic variants, which is higher in the hybrid mouse cell line than in human cells. This characteristic can explain the higher proportion of SNPs that could be phased in the mouse cell line compared with the human sample[2].

Finally, although the experimental methods for Hi-C are straightforward and relatively low cost, current protocols enabling genome assembly and haplotyping require 10^6–10^8 cells, a fairly large amount of material that may not be feasible to obtain, particularly from certain human patient samples.

As the new approaches[1–3] are developed further, it is likely that they will be even more effective if applied to sequence data from long-read technologies, such as those from Pacific Biosciences and Oxford Nanopore. It will also be interesting to see whether proximity ligation–based chromosome-scale scaffolding can facilitate the identification and characterization of genomic structural variations[16]. Eventually, a combination of these approaches[1–3] may enable the assembly of accurate human diploid genomes, enhancing the use of massively parallel sequencing for characterizing population structure, and ascertaining allelic contributions to gene expression and human diseases.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Burton, J.N. *et al. Nat. Biotechnol.* **31**, 1105–1112 (2013).
2. Selvaraj, S. *et al. Nat. Biotechnol.* **31**, 1113–1119 (2013).
3. Kaplan, N. & Dekker J. *Nat. Biotechnol.* **31**, 1139–1143 (2013).
4. Lieberman-Aiden, E. *et al. Science* **326**, 289–293 (2009).
5. Jin, F. *et al. Nature* **503**, 290–294 (2013).
6. The International Human Genome Sequencing Consortium. *Nature* **431**, 931–945 (2004).
7. Alkan, C., Sajjadian, S. & Eichler, E.E. *Nat. Methods* **8**, 61–65 (2011).
8. Gnerre, S. *et al. Proc. Natl. Acad. Sci. USA* **108**, 1513–1518 (2011).
9. The Genome 10K Community of Scientists. *J. Hered.* **100**, 659–674 (2009).
10. The 1000 Genomes Project Consortium. *Nature* **491**, 56–65 (2012).
11. Tewhey, R., Bansal, V., Torkamani, A., Topol, E.J. & Schork, N.J. *Nat. Rev. Genet.* **12**, 215–223 (2011).
12. Korbel, J.O. *et al. Science* **318**, 420–426 (2007).
13. Wheeler, D.A. *et al. Nature* **452**, 872–876 (2008).
14. The International Cancer Genome Consortium. *Nature* **464**, 993–998 (2010).
15. Simonis, M. *et al. Nat. Methods* **6**, 837–842 (2009).
16. Mills, R.E. *et al. Nature* **470**, 59–65 (2011).

Research Highlights

Papers from the literature selected by the Nature Biotechnology editors. (Follow us on Twitter, @NatureBiotech #nbtHighlight)

***Plasmodium falciparum* phosphoethanolamine methyltransferase is essential for malaria transmission**
Bobenchik, A.M. *et al. Proc. Natl. Acad. Sci. USA* **110**, 18262–18267 (2013)

Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing
Guo, H. *et al. Genome Res.* doi:10.1101/gr.161679.113 (31 October 2013)

Structural basis for the modular recognition of single-stranded RNA by PPR proteins
Yin, P. *et al. Nature* doi:10.1038/nature12651 (27 October 2013)

Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure
Ho, Y. *et al. Cell* **155**, 540–551 (2013)

Genetic identification of a neural circuit that suppresses appetite
Carter, M.E. *et al. Nature* **503**, 111–114 (2013)