

system provides an experimentally accessible testing ground for chiral ferromagnetic solids, where nanoscale skyrmions have enormous potential for technological applications and 3D hopfion nanoparticles could provide another dimension to exploit in data storage technology.

Finally, the aforementioned hopfion has linking number one, since any two field lines are linked exactly once. The integer, Q , which counts this linking number, is known as the Hopf invariant, and multi-hopfions are characterized by $Q > 1$. They can be constructed by twisting the skyrmion string through Q full revolutions before joining the ends. Multi-hopfions have not yet been

studied in chiral ferromagnetic media, but theories with hopfions that are stabilized by terms other than a chiral interaction show that it can be energetically favourable to tie a twisted skyrmion string in a knot rather than simply form a circular loop^{6,7}. Due to energetic constraints on the way the skyrmion can twist in a chiral medium, it is not clear whether stable multi-hopfions, or indeed knots, might exist in a chiral ferromagnetic host, but it is certainly an interesting avenue for future research. Frustrated magnets — where the formation of nanoscale skyrmions has been predicted⁸ and there are fewer constraints on the twist — may be the most promising

candidates to host nanoscale hopfions and nanoknots. □

Paul Sutcliffe is in the Department of Mathematical Sciences, Durham University, South Road, Durham DH1 3LE, UK.

e-mail: p.m.sutcliffe@durham.ac.uk

References

1. Manton, N. & Sutcliffe, P. *Topological Solitons* (Cambridge Univ. Press, 2004).
2. Yu, X. Z. *et al. Nature* **465**, 901–904 (2010).
3. Fert, A., Cros, V. & Sampaio, J. *Nat. Nanotech.* **8**, 152–156 (2013).
4. Ackerman, P. J. & Smalyukh, I. I. *Nat. Mater.* **16**, 426–432 (2017).
5. Evans, J. S., Ackerman, P. J., Broer, D. J., van de Lagemaat, J. & Smalyukh, I. I. *Phys. Rev. E* **87**, 032503 (2013).
6. Faddeev, L. & Niemi, A. J. *Nature* **387**, 58–61 (1997).
7. Batty, R. A. & Sutcliffe, P. M. *Phys. Rev. Lett.* **81**, 4798–4801 (1998).
8. Leonov, A. O. & Mostovoy, M. *Nat. Commun.* **6**, 8275 (2015).

GENE MEMORIES

The recent announcement by George Church and co-workers at Harvard University of plans to bring mammoths ‘back from extinction’ — more properly, to create elephants with mammoth features by splicing their genomes with fragments of those preserved in mammoth remains — has provoked some scepticism. But that optimistic scheme does at least remind us how good DNA is at preserving information over very long times.

That’s worth remembering if the notion of storing computer information in the chemical structure of a ‘fragile’ organic molecule seems risky. Church has also been one of the leading proponents of efforts to do just that: to use DNA as a storage medium for information technology. As over 20 years have passed since the notion of DNA computing was introduced¹, and as the capacity of DNA for encoding digital information was in fact apparent ever since Watson and Crick’s discovery of the double helix in 1953, it might seem surprising that it has taken so long for DNA storage to be seriously contemplated. What has made the difference is the rapid evolution over the past two decades of technologies for making long stretches of DNA to precise specifications at commercially viable cost, and likewise for reading out the information. Those techniques were not, of course, developed with data storage in mind, but to support gene manipulation and sequencing for biotechnology and genomics. They are augmented now with tools for

pinpoint editing of DNA sequences using the CRISPR/Cas9 system — giving would-be DNA data engineers a complete set of affordable tools for input, editing and readout.

The very real potential of DNA as a storage medium has been amply demonstrated. In 2012 Church *et al.* stored and read out a 5 MB file (a book) using microchips and advanced sequencing²; Goldman *et al.* subsequently showed how the approach might be scaled up while retaining 100% accuracy³. CRISPR/Cas9 has been used to find and edit specific stored strings in a DNA random-access memory⁴. All the same, encoding arbitrary data remains compromised by several factors. Whereas in principle each nucleotide (A, T, G and C) can encode two bits (00, 01, 10, 11), this coding capacity can’t be reached in practice. For one thing, some sequences — for example, those with high guanine-cytosine content — are hard to synthesize and read out accurately. Furthermore, variability in the synthesis and stability of some oligonucleotide sequences makes them unavailable for coding. So the actual storage capacity per nucleotide is limited to 1.8 bits. Previous studies have attained no more than 60% of this theoretical limit.

Erlich and Zielinski have now used an algorithm familiar to computer science, called fountain codes, to improve this performance⁵. It involves dividing the data set into short, non-overlapping segments in a way that screens and eliminates error-prone



PHILIP BALL

sequences. These message fragments, called droplets, are also tagged with random-number ‘barcode’ sequences for subsequent reassembly.

In this way Erlich and Zielinski could encode and read various data files — a computer operating system, an early motion picture, a computer virus and others — and not only read them back with perfect fidelity but also create error-free copies. The fountain algorithm achieves a capacity per nucleotide of 86% of the theoretical maximum.

This won’t yet make DNA data storage a commercial reality: it is still far too costly (about US\$3,500 per MB here). Yet not only are these costs still falling, but methods like this can also tolerate cheaper, less accurate synthesis methods than those developed for genomic technologies. □

References

1. Adleman, L. M. *Science* **266**, 1021–1024 (1994).
2. Church, G. M., Gao, Y. & Kosuri, S. *Science* **337**, 1628–1629 (2012).
3. Goldman, N. *et al. Nature* **494**, 77–80 (2013).
4. Tabatabaei Yazdi, S. M. H., Yuan, Y., Ma, J., Zhao, H. & Milenkovic, O. *Sci. Rep.* **5**, 14138 (2015).
5. Erlich, Y. & Zielinski, D. *Science* **355**, 950–954 (2017).