are also strongly magnetized[7], but none have been connected with intense electromagnetic flares.

The answer could relate to the fact that although magnetars are hotter at their surfaces than are other stars[5], they are fundamentally cryogenic objects. The outer layers of a neutron star contain heavy, neutron-rich nuclei, and begin to freeze into a solid soon after the stellar collapse that triggers the neutron star's formation[8]. This crust has peculiar properties when compared with those of Earth's outer layer. The temperature in the deep crust has dropped far below the melting point of the nuclear material from which it is made. Moreover, this material is an excellent conductor of electricity, effectively tying it to the twisting magnetic field — the magnetic field and crust must move together, either slowly during quiescence or more rapidly during an outburst.

Precisely how a magnetar flare is triggered is still being investigated. In contrast to the hot, magnetized atmosphere of the Sun, there are no swirling convective motions that actively deform the embedded magnetic field. But in a giant flare, we can be fairly certain that the crust undergoes a very large-scale disruption — imagine a tectonic event in which California and New York become interchanged. The slower component seen in the 15 April event and its siblings is consistent with what would be produced by the relaxation of a crustal deformation.

Such a disruption would twist up the exterior magnetic field of the magnetar, driving unstable electric currents one billion times stronger than those flowing through the Sun's corona[9,10]. A further consequence could be the ejection of a magnetic loop, similarly to large solar flares[11,12]. The magnetic disturbance would be strong enough to create a dense, outflowing gas of electrons, positrons and γ-rays. The interaction of such an electrically conducting gas with the magnetic field is thought to produce the subsecond-long γ-ray spectrum observed by Svinkin *et al.* and by Roberts and colleagues.

In its study[3], the Fermi LAT collaboration opens a new window on magnetar flares. It reports that the low-energy γ-rays described in the other two papers[1,2] were followed 19 seconds later by an emission of higher-energy γ-rays, which lasted for several minutes. This is the first detection of delayed, high-energy γ-rays from a magnetar flare. The proposed explanation involves the release of a cloud of fast-moving (relativistic) ions during the flare — the high-energy γ-rays are produced in a shock wave as the eruption hits the gaseous medium of NGC 253. However, it is unclear whether magnetar flares contain a substantial mass of ions. A pulse consisting of nearly pure electromagnetic radiation would also drive a shock wave, and might interact first with a relativistic nebula of particles confined around the magnetar.

Taking stock, the trio of papers reports a γ-ray flare that offers direct clues about how magnetic stresses relax in and around a neutron star. The measured energy fell short of that produced by most γ-ray bursts generated by colliding neutron stars, by a factor of 1,000 or more, although the duration was similar. Theorists modelling γ-ray bursts have so far failed to agree on the process that produces the events' signature γ-ray emissions. Understanding the similarities and differences with regard to magnetar flares will help to narrow down the possibilities. Continued monitoring of magnetars in nearby galaxies will also constrain models of the origin of fast radio bursts.

Christopher Thompson is at the Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario M5S 3H4, Canada.
e-mail: thompson@cita.utoronto.ca

1. Svinkin, D. *et al. Nature* **589**, 211–213 (2021).
2. Roberts, O. J. *et al. Nature* **589**, 207–210 (2021).
3. The Fermi LAT collaboration. *Nature Astron.* https://doi.org/10.1038/s41550-020-01287-8 (2021).
4. Weltman, A. & Walters, A. *Nature* **587**, 43–44 (2020).
5. Kaspi, V. M. & Beloborodov, A. M. *Annu. Rev. Astron. Astrophys.* **55**, 261–301 (2017).
6. Hurley, K. *et al. Nature* **434**, 1098–1103 (2005).
7. Ferrario, L., de Martino, D. & Gänsicke, B. T. *Space Sci. Rev.* **191**, 111–169 (2015).
8. Ruderman, M. A. *Nature* **218**, 1128–1129 (1968).
9. Thompson, C. & Duncan, R. C. *Astrophys J.* **561**, 980–1005 (2001).
10. Thompson, C., Yang, H. & Ortiz, N. *Astrophys. J.* **841**, 54 (2017).
11. Lyutikov, M. *Mon. Not. R. Astron. Soc.* **367**, 1594–1602 (2006).
12. Parfrey, K., Beloborodov, A. M. & Hui, L. *Astrophys. J.* **774**, 92 (2013).

Genetics

Repeat DNA expands our understanding of autism

Anthony J. Hannan

A link has been found between repetitive stretches of DNA called tandem repeats and autism spectrum disorder. The discovery might inform approaches to studying tandem repeats in a wide range of other human disorders. **See p.246**

Approximately half of the human genome, known as the repeatome, consists of repetitive DNA sequences. The repeatome includes more than one million tandem repeats — sections of DNA in which a sequence is replicated many times in tandem — whose biology remains largely unexplored. More than 50 diseases are known to be caused by expansion of a tandem-repeat sequence in a single gene; among them are Huntington's disease and fragile X syndrome[1]. But less-well understood is the role of tandem repeats in polygenic diseases, which have more-complex genetic underpinnings. On page 246, Mitra *et al.*[2] use a newly developed bioinformatics approach to identify tandem repeats associated with one such condition, autism spectrum disorder.

Autism spectrum disorder (ASD) is highly prevalent, affecting approximately 1–2% of children in the United States (see go.nature.com/38sqvhd), although this varies internationally. It is characterized by atypical neurodevelopment, communication deficits, atypical social functioning, restricted interests and repetitive behaviours. Although progress is being made[3] in discovering the genetic basis of ASD, it remains poorly understood. Changes in the number of copies of large segments of DNA, along with other genetic variants, have been previously implicated[3], but the capacity to systematically investigate tandem repeats genome-wide has been optimized only in the past few years, thanks to advances in DNA sequencing and bioinformatics[1].

Mitra *et al.* analysed tandem repeats in people who have ASD, and in their immediate families, using a newly developed bioinformatics tool that they named MonSTR. The tool processes DNA sequences from tandem-repeat regions across the genome to determine the likelihood that a *de novo* mutation (one that occurs only in the person who has ASD, and not in their parents) led to a change in a tandem repeat.

The analysis revealed that tandem-repeat mutations are significantly more common in people who have ASD than in their unaffected siblings, with mutations more likely to cause the repeat to expand than to contract (Fig. 1). Many of the expansions occurred in DNA regions that drive the expression of genes involved in fetal brain development. To determine the likelihood that a mutation would be deleterious, Mitra *et al.* developed another

bioinformatics tool (named SISTR) based on an evolutionary model of tandem-repeat variation. This analysis revealed 25 mutations, present in individuals with ASD, that were likely to be the most harmful; these mutations are rare in the general population, presumably because they are strongly selected against. Furthermore, some of these tandem-repeat mutations occurred in genes that had previously been associated with ASD, increasing the likelihood that they are directly involved in the disorder.

Mitra and colleagues' work echoes a recent study by Trost and colleagues[4], which examined tandem repeats in ASD using a different, complementary approach. Trost *et al.* compared tandem-repeat lengths in thousands of genomes in people who had ASD and those who did not (some family members, and some unrelated), using another recently developed bioinformatics tool. This approach allowed the authors to study more cases and controls than Mitra *et al.* (approximately 17,000 genomes, compared with approximately 6,500 genomes in total). The tool used by Trost and colleagues, named ExpansionHunter Denovo[5], is optimized for detecting longer repeats (at least 150 bases pairs of DNA) and their expansions, whereas MonSTR also allows analysis of relatively short tandem repeats, and of both expansions and contractions.

As Mitra and colleagues have done, Trost *et al.* identified many tandem-repeat mutations associated with autism. But because the two studies used quite different approaches, the identified mutations were largely complementary, rather than overlapping. Another difference is that Trost and colleagues cannot say exactly how many of the mutations are inherited, rather than *de novo*, because many of the people in their control groups were unrelated to the people with ASD. And Trost *et al.* found specific associations between tandem-repeat expansions and particular clinical features associated with cognitive function, notably lower IQ and adaptive ability, which Mitra and colleagues did not investigate. Analysis of these clinical subtypes will be valuable in future studies of tandem repeats in ASD.

The large number of non-overlapping tandem-repeat mutations identified in these two papers together provide compelling evidence that tandem repeats contribute significantly to the genetic burden associated with ASD. These studies will inform our understanding of the mechanisms that underlie the condition, as well as future approaches to diagnosis and treatment. However, ASD is a complex and varied disorder. Although both studies involved thousands of individuals, larger international replication studies are needed to establish how robust these genetic associations are, and how commonly each mutation is associated with ASD in general, as well as with each clinical subtype.

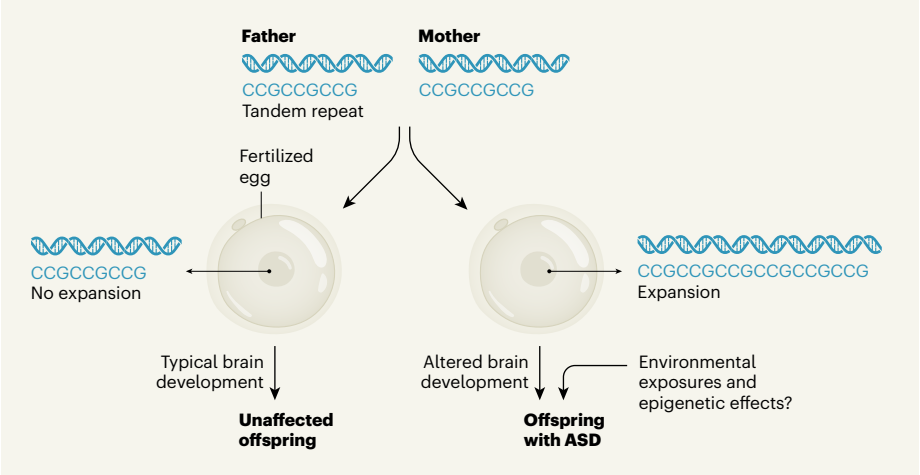

Figure 1 | Autism-linked mutations in families. Tandem repeats are DNA sequences repeated many times in tandem (here, an example repeat involves nucleotides dubbed CCG). Although present in all individuals, tandem repeats can cause disease if they expand. Mitra *et al.*[2] developed a bioinformatics tool to identify tandem repeats that were mutated from their normal length, and applied it to families that include one offspring who has autism spectrum disorder (ASD). Tandem-repeat expansions in people who have ASD (which presumably arose in the fertilized egg) often occurred in genetic regions where they might lead to altered brain development, whereas such expansions were less common in unaffected siblings. It remains to be determined whether and how these repeat expansions — along with other genetic factors, environmental factors and epigenetic effects (molecular changes that alter gene expression without modifying DNA sequence) — affect the brain to predispose an individual to developing ASD.

In addition, both studies are correlative. There is not yet any direct evidence that the identified mutations have a role in disease, except in the case of a small subset identified by Trost and colleagues. This consists of mutations that occur in genes whose repeat expansions are known to cause other tandem-repeat disorders[1], including fragile X syndrome, myotonic dystrophy and Friedreich ataxia. Therefore, an urgent priority now is to test the role of all of the identified tandem repeats in animal models, and in human stem-cell cultures derived from people who have ASD and their families. These further studies will enable key questions to be addressed. Does each tandem repeat regulate aspects of brain development and function (as well as peripheral systems implicated in ASD)? If so, what mechanisms are involved? And are specific tandem repeats associated with subtypes of ASD or particular co-morbidities, given that the condition often occurs with one or more other disorders, including epilepsy, intellectual disability and attention deficit hyperactivity disorder?

Many other avenues for future research are also apparent. For example, do some tandem-repeat lengths confer an evolutionary advantage? If so, perhaps ASD associated with tandem-repeat mutations results from an evolved mechanism that makes tandem-repeat length more variable (a form of genetic plasticity) than that of non-repetitive sequences[1]. For a given gene, although certain tandem-repeat lengths might have selective advantages, it is possible that extremely long or short repeats (and perhaps, in particular environmental contexts, repeats of a particular length) could be associated with human disorders.

Another question to be addressed is whether each tandem-repeat mutation is differentially detrimental on different genomic and environmental backgrounds, causing or promoting ASD in some settings but not in others. We know that many tandem repeats are highly responsive to environmental signals that lead to altered epigenetic modifications (molecular marks on DNA that can alter gene expression without changing the underlying DNA sequence). Indeed, some tandem repeats might themselves act as regulators of these modifications[1].

More broadly, the repeatome — and tandem repeats, in particular — should now be systematically studied across a range of common human disorders, including cancer, diabetes and brain disorders such as schizophrenia and depression. Genome-wide association studies (GWAS), which link genetic variants to traits and disorders of interest, have improved our understanding of these polygenic disorders, but substantial gaps remain[6]. It has been proposed that such 'missing heritability' might be explained, at least in part, by tandem repeats[7]. But many of these are highly variable, and so are unlikely to be discovered by GWAS approaches, which map genes on the basis of variation at single nucleotides[7]. The approaches outlined in the current papers[2,4] and elsewhere[1,6,8,9] provide a road map for identifying some of this missing heritability across a broad spectrum of human traits and disorders.

In the long term, taking a comprehensive approach to analysing and understanding tandem repeats (and the repeatome more generally) could enable mutations to be corrected in the clinic, using therapeutic tools such as CRISPR–Cas gene-editing techniques and drugs[10]. It is possible that such approaches, as well as other future therapies developed thanks to our expanded understanding of repeatome biology, will constitute a fundamental foundation of precision medicine in the twenty-first century.

Anthony J. Hannan is at the Florey Institute of Neuroscience and Mental Health, University of Melbourne, Parkville, Victoria 3052, Australia, and in the Department of Anatomy and Neuroscience, University of Melbourne.
e-mail: anthony.hannan@florey.edu.au

1. Hannan, A. J. *Nature Rev. Genet.* **19**, 286–298 (2018).
2. Mitra, I. *et al. Nature* **589**, 246–250 (2021).
3. Schaaf, C. P. *et al. Nature Rev. Genet.* **21**, 367–376 (2020).
4. Trost, B. *et al. Nature* **586**, 80–86 (2020).
5. Dolzhenko, E. *et al. Genome Biol.* **21**, 102 (2020).
6. Eichler, E. E. *et al. Nature Rev. Genet.* **11**, 446–450 (2010).
7. Hannan, A. J. *Trends Genet.* **26**, 59–65 (2010).
8. Breuss, M. W. *et al. Nature Med.* **26**, 143–150 (2020).
9. Mousavi, N. *et al. Bioinformatics* https://doi.org/10.1093/bioinformatics/btaa736 (2020).
10. Nakamori, M. *et al. Nature Genet.* **52**, 146–159 (2020).

Psychology

Debate about universal facial expressions goes big

Lisa Feldman Barrett

An analysis of more than 6 million YouTube videos finds that people around the world make similar facial expressions in similar social contexts. The study brings data science to the debate about the universality of emotion categories. **See p.251**

When you are angry, do you scowl, cry or even laugh? To what extent do your facial movements depend on the situation you are in — whether you are in a formal meeting, say, or at home with your family? And do other people around the world express anger in such situations in the same way? These questions are at the centre of a contentious scientific debate about the nature of emotion that has raged for more than a century. On page 251, Cowen *et al.*[1] enter the fray. The authors used a type of machine learning to investigate whether people adopt the same facial expressions in similar contexts across cultures. Their results are sure to be the subject of lively discussion.

The universality debate is central to an understanding of the nature, causes and functions of emotions. The universality hypothesis proposes that people around the world consistently use certain configurations of facial-muscle movements to specifically express instances of a certain category of emotion. For example, people are said to frown in sadness consistently enough (and frown infrequently at other times) for frowning to be recognized as the expression of sadness the world over. The same goes for scowling in anger, smiling in happiness and so on. Such expressions are thought to have evolved to signal emotional information in contexts that posed fitness challenges for our hunter-gatherer ancestors[2,3]. Sometime in our evolutionary past, the hypothesis goes, our ancestors capitalized on specific, universal emotional expressions as a means of communicating those challenges to one another, helping them to survive and reproduce.

There is considerable debate about whether the published empirical evidence provides support for the universality hypothesis. A large body of literature seems to back it up, but these studies have been consistently criticized for methodological problems[4,5]. Furthermore, a growing number of experiments (both in modern, urban populations[5] and in small-scale foraging communities[6], such as the Hadza hunter-gatherer group in Tanzania[7]) using diverse methods seem to call the universality hypothesis into doubt. This research suggests that expressions of emotion might be context-specific, and therefore variable across instances — the facial movements in each instance being tailored to the immediate context, as is the case for all other motor movements. In fact, study after study has revealed the potent influence of context on both how people move their faces to express emotion and how other people infer emotional meaning in those facial movements[5,8,9].

> **"Perhaps just one other study has so far matched this broad scale of sampling."**

For some scientists, these observations seem consistent with alternative evolutionary hypotheses of emotion[5], but for others, the findings introduce another question about universality. Might situated emotional expressions — how people express emotion in certain situations — be universal across cultures? This is the question that Cowen *et al.* set out to answer, by investigating whether certain facial configurations are found in similar contexts worldwide.

Previous studies have almost exclusively used artificial methods to observe how people express emotion with facial movements and infer emotional meaning in facial configurations. One common task is to arm people with a small, preselected set of emotion words (such as 'anger' or 'sadness') and to ask them to label posed, disembodied, contextless faces (such as a person smiling) with the word that they think best describes the emotion on each face. This method, when compared with others, has been consistently shown to inflate support for the universality hypothesis[4–6,10,11]. A related method offers people a single, impoverished scenario for an emotion category ('You have been insulted, and you are very angry about it', for instance). They are then asked to pose the facial configuration they believe they would make to express instances of that emotion category. Both approaches seem to encourage people to rely on stereotypes about emotional expressions that do not typically reflect the varied ways in which people express emotion in everyday life[5].

One major strength of Cowen and colleagues' effort is that they analysed facial configurations in more-natural settings. The authors curated YouTube videos of people in natural social contexts, such as at weddings, sports events or playing with toys (Fig. 1). Another strength is the scope of their sampling: the paper sampled more than 6 million videos from 144 countries in 12 regions around the world. Perhaps just one other study[12] has so far matched this broad scale of sampling.

Cowen and colleagues used a powerful machine-learning method involving what are called deep neural networks (DNNs) to assess the extent to which specific facial configurations (called facial expressions by the authors) could be reliably observed in the videos across cultures. They trained one DNN to classify the facial configurations in the videos as emotional expressions and a second DNN to classify the videos' contextual elements. Then they estimated the associations between the two classifications to determine how frequently each class of facial expression occurred in videos that were classified as containing similar contextual elements, and compared