Minireview

Variable number tandem repeats – Their emerging role in sickness and health

Jack NG Marshall[1], Ana Illera Lopez[1], Abigail L Pfaff[2,3], Sulev Koks[2,3], John P Quinn[1] and Vivien J Bubb[1]

[1]Department of Pharmacology and Therapeutics, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool L69 3BX, UK; [2]Perron Institute for Neurological and Translational Science, Perth, WA 6009, Australia; [3]Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Perth, WA 6150, Australia
Corresponding author: Vivien J Bubb. Email: jillbubb@liverpool.ac.uk

Impact statement

Interpretation of the functional consequences of human genome variation is an essential part of modern biomedical research. This review describes the impact of repetitive DNA on experimental biology by allowing insight into the role of non-coding DNA in regulation of gene structure and function. The genetic variation inherent in repetitive DNA is becoming an increasing important parameter in personalized medicine. Emerging long read DNA sequencing technologies should aid the improved characterization of these elements in the human genome which will integrate that knowledge into clinical practice and improve the precision of clinical diagnoses and decision-making. This review should prove a valuable resource to the field to capture the distinct mechanism utilized by repetitive DNA in gene function.

Abstract

Understanding the mechanisms regulating tissue specific and stimulus inducible regulation is at the heart of understanding human biology and how this translates to wellbeing, the ageing process, and disease progression. Polymorphic DNA variation is superimposed as an extra layer of complexity in such processes which underpin our individuality and are the focus of personalized medicine. This review focuses on the role and action of repetitive DNA, specifically variable number tandem repeats and SINE-VNTR-*Alu* domains, highlighting their role in modification of gene structure and gene expression in addition to their polymorphic nature being a genetic modifier of disease risk and progression. Although the literature focuses on their role in disease, it illustrates their potential to be major contributors to normal physiological function. To date, these elements have been underreported in genomic analysis due to the difficulties in their characterization with short read DNA sequencing methods. However, recent advances in long read sequencing methods should resolve these problems allowing for a greater understanding of their contribution to a host of genomic and functional mechanisms underlying physiology and disease.

Keywords: VNTR, SVA, transcriptional regulation, neurological disease, physiology

Experimental Biology and Medicine **2021; 246: 1368–1376. DOI: 10.1177/15353702211003511**

Introduction

Ninety-eight percent of the human genome does not code for proteins, but contains a variety of regulatory elements including those to direct (a) tissue specific and stimulus inducible gene expression, (b) differential mRNA splicing to generate distinct protein profiles or turnover, and (c) functional non-coding RNAs. The properties of these regulatory components like the exons themselves can be modified by genetic polymorphism or mutations. To date, genome analyses have favored short read next generation DNA sequencing and genome-wide association studies (GWAS). The latter has focused on analyzing variation in single nucleotide polymorphisms (SNPs) and has identified genomic variation in strong association with phenotype both in exons and non-coding regions. Such SNPs in exons are often easy to interpret mechanistically as many will change a key amino acid providing an explanation for the phenotypic association. However, the functional significance of the vast majority of SNPs in non-coding DNA is difficult to determine; furthermore, a major shortfall of this approach is that it has neglected the role and function of other sources of genetic variation in the genome such as that represented by repetitive DNA. This is in part due to the inherent instability of human repetitive DNA in *E. coli* cloning strategies used in the construction of the library of DNA sequences incorporated into the reference genome, and the difficulties in interpreting and reassembling sequence from such short read sequences in areas of repetitive DNA. The advent of long-read sequencing technologies will help overcome the latter problem, and technologies such

as Oxford Nanopore or Pacific Biosciences single-molecule real-time sequencing (PacBio) can easily read 10 kb or more sequence; furthermore, base calling accuracy has improved (error rate said to be <1% for PacBio and <5% for Oxford Nanopore) and analytical pipelines to call structural variation have been developed.[1] This review will focus on one class of repetitive DNA, namely variable number tandem repeats (VNTRs), highlighting their functional roles in genome regulation and as biomarkers of disease focusing on neurological diseases and disorders

Consequences and genetic nature of tandem repeat DNA

Tandem repeat (TR) DNA imparts a huge source of variation to the genome, and TRs have evolved throughout evolution and contribute to genetic diversity. Although they are found in many eukaryote species, present day interest has focused on their role in hominid evolution and their contribution to the development of human-specific traits in the evolution of modern humans and in particular neural function.[2–4] In modern humans, TR DNA has been associated both with traits such as aggression and addictive behaviors, reviewed by literature[5–8] and implicated in the increased susceptibility to, and risk of developing, a wide variety of neurological diseases such as motor neuron disease[9–11] and Alzheimer's disease [12,13] mental health conditions such as bipolar disorder, depression, and schizophrenia.[14,15] To date, 1584 such repeats have been found to be human specific and it has been proposed that they contribute to human-specific traits.[4] When the repeat number is variable, i.e. polymorphic, they become VNTRs, a class which accounts for approximately 3% of the human genome[16,17]; such repeat elements can vary in copy number and additionally demonstrate single nucleotide variation within the repeats or even short insertions or deletions. There are several mechanisms proposed, by which this variation could develop, including errors in DNA replication such as slipped-strand mispairing which results in the misalignment of DNA strands and thus expansion or contraction of the copy number of the DNA motifs,[18] homologous recombination, and duplications. Recently it has been shown 55% of VNTRs map to the terminal 5Mbp of human chromosomes,[19,20] and regions that have previously been found to demonstrate increased double-strand breaks during early stages of meiosis[21,22] and male meiotic recombination. VNTRs may be located within exons, introns, or intergenic spaces and this aforementioned group of VNTRs appears to favor intronic locations.[4] A second pathway contributing to the derivation of VNTRs is via retrotransposition, particularly utilizing the composite SINE-VNTR-*Alu* (SVA) element, and these elements tend to show bias against incorporation in to genic regions, and have a high GC content.[4] Unsurprisingly, VNTRs located within exons are the smallest group. Originally the main focus of research was on short tandem repeats (STR) (or micro-satellites) as it became apparent that expansions of these can occur in somatic cells, and this phenomenon is implicated in diseases such as fragile X-associated tremor/ataxia syndrome, myotonic dystrophy, Huntington's disease, spinocerebellar ataxia, and amyotrophic lateral sclerosis (ALS).[23–33] This approach originally identified a number of diseases associated with triplet expansions, i.e. Fragile X syndrome in 1991 [23] and Huntington's disease in 1993;[34] however, expansions of longer repeat sequences have since been found. For the latter, examples include: pentanucleotide expansions in the genes *SAMD12* and *RFC1* found in benign adult familial myoclonal epilepsy type1 and ataxia syndromes, respectively,[35,36] a hexanucleotide repeat expansion in the *C9ORF72* gene associated with ALS,[32,33] and a 12-nucleotide expansion in the *CSTB* gene associated with progressive myoclonic epilepsy of the Unverricht-Lundberg type or EPM1 [37] (Table 1). To date, incidences of personal somatic expansions or generation-to-generation instability in TRs of much larger repeat length have not been identified other than the intriguing recent report of an expanded VNTR sequence (around 25 nucleotides in length) in an intron of the *ABCA7* gene, and the repeat number ranges from 12 to 427 or greater, and has been identified as a risk factor in Alzheimer's disease; however, it has not been reported whether the repeat number seen in this expansion is stable between germ and somatic cells from the same individual.[13] In general, based on current sequence information, the polymorphic but stable expansions seen in longer TRs appear to pre-date the evolution of modern human populations and have been used to track ancient population migrations. Furthermore, for some VNTRs, the allele frequencies or base pair composition of the repeat unit have been found to differ between populations and ethnic groups throughout the world including *TRIB3,*[47] *WRD7,*[42] and *DNAJC5/miR-941.*[48] This is a factor which must be taken into account when assigning "risk" to such alleles as exemplified by the hexanucleotide expansion associated with ALS in the *C9ORF72* gene which is one of the four most common causes of ALS in Caucasian populations but is much rarer in Chinese and other East Asian populations.[39,49]

Assessing the functional potential of VNTRs as transcriptional regulators of gene expression

An increasing number of VNTRs have been identified which support differential gene expression both *in vivo and in vitro.* Perhaps one of the earliest and most striking illustrations of allele-specific activity, was the *in vivo* demonstration of the ability of the human-specific *SLC6A4* intron 2 VNTR to direct differential reporter gene expression in the midbrain of mouse embryos equivalent to the area where the mouse serotonin transporter is initially expressed in the developing brain.[43,50] Association studies have established a link between repeat copy number in this VNTR as a risk factor in various neurological diseases or disorders; however, establishing the functional significance of such VNTR polymorphism *in situ* is more problematic. Some of this difficulty may be attributed to non-coding VNTRs being regulatory domains that are only functional in specific tissues, developmental stages, or in response to specific cellular challenges. Recent examples of VNTRs found to have regulatory properties include the finding

Table 1. Examples from the text of neurological diseases and disorders which have been associated with polymorphic tandem repeats.

Gene	STR/VNTR/SVA	Disease association	Refs
FMR1	STR (triplet)	Fragile X-associated tremor/ataxia syndrome	23,24
DMPK	STR (triplet)	Myotonic dystrophy type 1	28,29
HTT	STR (triplet)	Huntington's disease	30,34
ATXN1	STR (triplet)	Spinocerebellar ataxia type 1 amyotrophic lateral sclerosis	11,31
NIPA1	STR (triplet)	Hereditary spastic paraplegia type 6, amyotrophic lateral sclerosis	9
NOTCH2NLC	STR (triplet)	Neuronal intranuclear inclusion disease-related disorders, Parkinson's disease	38
SAMD12	VNTR (pentamer)	Benign adult familial myoclonal epilepsy type1	35
RFC1	VNTR (pentamer)	Ataxia syndrome	36
C9ORF72	VNTR (hexamer)	Amyotrophic lateral sclerosis frontotemporal dementia	10,32,33,39
CSTB	VNTR (12 mer)	Progressive myoclonic epilepsy of the Unverricht-Lundberg type or EPM1	37
ABCA7	VNTR (approx.25 mer)	Alzheimer's disease	13
SLC6A3	VNTR (40 mer)	Attention-deficit hyperactivity disorder	40,41
DRD4	VNTR (48 mer)	Attention deficit hyperactivity disorder, addictive, and eating disorders	5,8
WRD7	VNTR (69 mer)	Amyotrophic lateral sclerosis	42
SLC6A4	VNTR (22–23 mer and 16–17 mer)	Anxiety-related traits Affective disorders	7,14,43
MAOA	VNTR (30 mer and decamer)	Impulsive/anti-social/aggressive behaviors, affective disorders	6,44,45
TAF1	SVA	X-linked dystonia Parkinsonism	46

that the longer risk alleles of the intronic VNTR in *ABCA7* favored use of a cryptic splice site resulting in exon skipping [12] and that the number of VNTR repeats in the promoter region of *TRIB3* correlated with mRNA levels in various tissues.[47]

If VNTRs are involved in transcriptional or post transcriptional gene regulation then they will act in consort with other gene regulatory domains to direct tissue specific and stimulus inducible gene expression. For example, it has been shown that a VNTR located at the *MIR137* locus works together with the SNP rs2660304 to drive differential promoter activity *in vitro*, thereafter it was shown using haplotype analysis at the *MIR137* locus that rs2660304 was a proxy SNP for the schizophrenia GWAS SNP rs1625579.[51] Furthermore, interactions between multiple VNTRs to regulate gene expression are possible. *In vitro* analysis has demonstrated that distinct VNTRs within a gene locus can act either independently or synergistically to regulate transcription. For example, the serotonin transporter (*SLC6A4*) gene which has VNTRs located both in the linked polymorphic region of the 5′ promoter and intron 2 which have been shown to act combinatorically,[52] a further example is the *MAOA* gene which in addition to the well-characterized μVNTR has a second distal (d)VNTR[44] located approximately 500 bp upstream of the μVNTR.[53] The complexity this produces was illustrated by the recent meta-analysis of *MAOA* by Tunbridge *et al*. where it was found that the high and low activity μ alleles were associated with enzyme activity in the blood but did not affect MAOA mRNA abundance;[45] concurrently, Manca *et al*. demonstrated that the distal VNTR regulated mRNA levels of the canonical isoform of MAOA in a cell line model.[54] Moreover, specific regulatory effects on MAOA isoforms and an additive effect of the two VNTRs at this locus were highlighted in additional work.[53] A further layer of intricacy was uncovered by the finding that VNTR (allele) specific responses to a stimulus also existed.[54] Thus, it is important to consider the overall haplotype when analyzing gene regulation rather than assuming a single regulatory variation is solely associated with a specific phenotype.[47,53]

Proposed mechanisms of action and emerging novel roles for VNTRs to alter the genomic transcriptome

In addition to the more conventional mechanism of provision of transcription factor (TF) binding domains by which VNTRs can contribute to the regulation of transcription, there are several examples of them encoding miRNAs and affecting epigenetic parameters (Figure 1). In general, VNTRs have been shown to function as transcriptional regulatory domains, by providing binding sites for transcription factors and modulating the affinity of binding on the basis of repeat unit copy number.[55–57] More specifically in the context of neurological disorders, the previously mentioned *SLC6A4* intron 2 VNTR has been assessed via reporter gene assays in rat prefrontal cortical cells, where it was found that the different copy number variants induced reduced but differential reporter gene expression in

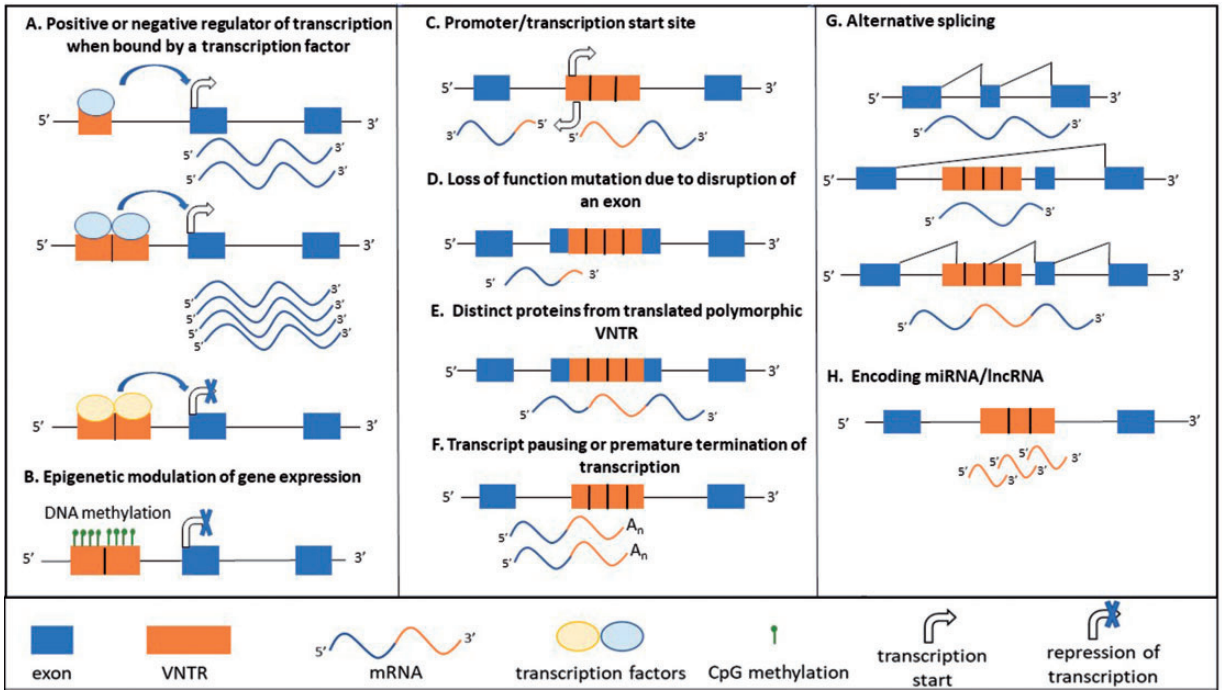

Figure 1. Mechanisms by which tandem repeat DNA can modify gene expression. Differential regulation at an allele can result either as a result of polymorphic repeat number, SNPs or indels in the repeats, pathogenic expansion of the repeat or in the case of SVAs presence or absence polymorphism. (A color version of this figure is available in the online journal.)

response to CCCTC-binding factor (CTCF).[52] Furthermore, it has been shown that both *SLC6A4* VNTRs can bind multiple transcription factors, including y-box binding protein (YB1), CTCF and methyl-CpG binding protein (MeCP2), inducing allele-specific expression profiles in response to cocaine.[58] A further example is the VNTR identified in the 3′ UTR of the human dopamine transporter (*SLC6A3*) gene, which has been associated with attention-deficit hyperactivity disorder.[40,41] Previously it had been shown using reporter gene assays in the SH-SY5Y cell line that the *SLC6A3* VNTR induced significant repression of luciferase expression in response to the HESR (HEY) family of transcription factors, specifically HESR1 and HESR2.[59]

More recently, it has been discovered that some VNTRs have the potential to encode for miRNAs. This phenomenon has been recently described for VNTRs within the genes *WDR7* [42] and *DNAJC5* which harbors the human-specific miR-941 within an intronic VNTR.[48] This latter example displays unusual features for a recently emerged human-specific miRNA, namely high levels of expression, particularly of note with regard to this review, in the cerebellum and prefrontal cortex and the copy number demonstrates a high level of variability. Human-specific regulation by this miRNA has been shown in the brain, targets included the host gene *DNAJC5* whose protein has amongst other functions a role in neurotransmitter release,[48] preventing neurodegeneration [60] and is associated with adult-onset neuronal ceroid lipofuscinosis.[61] Furthermore, this poses the question does the copy number of the VNTR leads to distinct levels of the miRNA directly, or is the mechanism indirect, achieved by modification of the processing of the internal miRNA? The putative miRNA encoded within *WDR7* (incidentally itself a target of miR-941[48]) has been detected in cytoplasmic aggregates or speckles when experimentally overexpressed. Such RNA foci, albeit in these cases nuclear, are recognized as important features in a number of RNA gain of function disease models such as that for myotonic dystrophy [62] and *C9ORF72*-associated ALS.[63,64]

Epigenetic modification of a VNTR could have significant consequences for all VNTR-directed regulatory mechanisms in both the immediate response to cellular signalling and the medium- and long-term properties of the VNTR to modulate gene function. This was illustrated by the analysis undertaken by Vasiliou *et al.* regarding the response of the *SLC6A4* serotonin transporter gene to stimulation by cocaine—differential effects on transcription factor binding were seen dependant on which VNTR allele was being examined and correlated with MeCP2 binding.[58] One clear mechanism is methylation of the VNTR itself, this requires the VNTR to contain CpG targets. Simplistically, the more CpG in the VNTR, the greater opportunity for methylation changes to alter VNTR function. This is perhaps obvious in short repeats of CG, but can also occur in expansion repeats such as the hexamer CCCCGG tandem repeat found in *C9ORF72*.[65,66] Changes in the status of CpG methylation have long been noted to affect the stability of repetitive elements containing such sequences; furthermore, methylation status has been purported to enable DNA to adopt alternative non-B DNA structural forms such as Z-DNA.[67] Consideration of the methylation status of the VNTR may help explain some of the conflicting reports of the association of repetitive elements with traits and disease. A prime example of this is the *MAOA* gene reviewed by Ziegler and Domschke, and this highlighted that the μVNTR lies in an area that exhibits

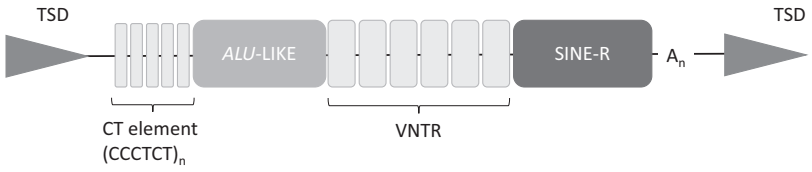

Figure 2. Illustration of the consensus structure of the non-long terminal repeat retrotransposon SVA element (~0.7–4 kb). The VNTR and CT elements can both be polymorphic for the number of tandem repeats and individual single repeat elements can also be polymorphic with SNPs and/or indels, and the polyA (A_n) may also be polymorphic in length. The element is flanked by target site duplications (TSD).

differential methylation; furthermore, the level of methylation is postulated to contribute to the different disease profiles that have been attributed to the *MAOA* gene, and the location of *MAOA* on the X chromosome could also contribute to the gender differences seen in the various analyses.[68] This observation was expanded by Manca *et al.*, who demonstrated in a cell line model that there was allele-specific gene expression and transcription factor binding with corresponding allele-specific epigenetic marks in the region encompassing the μVNTR; furthermore, exposure to the mood stabilizer sodium valproate resulted in an allele-specific response in transcription factor binding and epigenetic marks.[54]

The final mechanism to be highlighted is modification of genome structure and in particular formation of G-quadruplexes (G4). These structures can be generated by repetition of C-rich hexameric sequences such as CCCTCC and CCCCGG and variants thereof. The classical model for G4 structure is the trimeric repeat of CCCTCCCC found in the *MYC* gene promoter which has been associated with promoter usage for MYC expression.[69,70] Although this example is not a VNTR, it has laid the basis for the putative action of such repeats found in VNTRs in mechanisms involved in disease progression or initiation as exemplified by the CCCCGG expansion repeat in the *C9ORF72* gene associated with ALS.[71] Such repeats are postulated to form a genomic structure that allows for torsional strain on one side and opening of the genome on the other thus allowing for a range of distinct structural parameters to be imposed at local regions of the genome. A number of transcription factors, both double stranded DNA binding factors such as CTCF and Sp1, and single stranded nucleic acid binding factors, such as hnRNPK and YB1, could then bind to such VNTRs to determine localized genomic structure.

VNTRs in non-long terminal repeat retrotransposons

VNTRs are often considered as standalone elements in the genome but as discussed earlier they are also key domains of larger functional components such as the compound SINE-VNTR-*Alu* (SVA) element, a non-long terminal repeat retrotransposon. Although the total number of SVAs in the human reference genome is modest (approx. 2700), SVAs are hominid-specific elements with almost 50% of those characterized in the reference genome found to be human specific and have been postulated to correlate with the development of hominid lineage-specific traits.[2] They are composite regulatory DNA domains containing multiple regulatory components (Figure 2) that affect reporter gene expression.[72] Several of these domains are tandem repeats, specifically the CT hexamer flanking domain and a central large VNTR domain. The VNTR component of the SVA is variable in length, and it can consist of a single VNTR or two separate VNTRs (which are not necessarily both variable in number) which may share related primary sequence but are clearly distinct from one another. The VNTR may also contain more than one specific repeat motif. SVAs also have the potential to form the structural components outlined above, namely G-quadruplexes, and this ability is embedded in the CT hexamer and the central VNTR region where the CpG content can approach 60%.

The association of SVAs with disease may not only be a modifying parameter but can also be causative of the disease. This supports the main drive of our hypothesis that VNTRs are not only biomarkers of disease but can also be mechanistically involved in progression of the disease. This is demonstrated by the observation that some SVAs can be polymorphic for their presence or absence in the genome, thus for specific SVAs this means that they can be (1) associated with methylation differences at adjacent promoters as exemplified by the *LRIG2* gene,[73] (2) associated with differential RNA expression and disease severity in Parkinson's disease (Pfaff *et al.* in preparation), and (3) causative of disease as seen in X chromosome-linked dystonia Parkinsonism.[46]

Summary

This review highlights the emerging field analyzing the role and importance of VNTRs in regulation of genome function and regulation of the cell transcriptome. These elements have the potential to function in many different ways and have been postulated to contribute to human evolution. It is apparent, however, that each individual element should not be considered alone, rather it is their concerted action that is important, and this provides a layer of intricacy to regulation allowing an exquisite and specific response to a stimulus. In the recent past, VNTRs have also been found to be linked with many neurological conditions and disorders and their activity is expected to explain some of the missing heritability in such disorders. To date, analysis of VNTRs in association studies has often relied on labor intensive techniques such as polymerase chain reaction (PCR), which itself can be technically challenging in these domains due to their repetitive nature and in some instances, high GC content. However, with the advent and increased availability of long-read sequencing methods, these problems will in part be overcome and allow a more rapid and robust

analysis thus aiding the identification of a host of genomic and functional mechanisms underlying our physiology. For example, long-read sequencing has recently enabled the identification of a short tandem repeat (GGC) expansion in the *NOTCH2NLC* locus which can have as many as 517 repeats in affected individuals with neuronal intranuclear inclusion disease-related disorders[38] and facilitated epigenomic profiling of transposable elements.[74]

AUTHORS' CONTRIBUTIONS

JNGM, JPQ, and VJB conceived the idea, all authors reviewed the literature, wrote, and edited the article.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article

FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: JNGM was funded by a Medical Research Council doctoral training studentship [grant number MR/N013840/1]. AIL was funded by a Wellcome Trust PhD Studentship [grant number 109095/Z/15/Z]. ALP and SK are funded by MSWA, The Michael J. Fox Foundation, Shake It Up Australia and The Perron Institute. JPQ and VJB are funded by the Andrzej Wlodarski Memorial Research Fund.

ORCID iDs

Abigail L Pfaff https://orcid.org/0000-0002-2231-9800
Sulev Koks https://orcid.org/0000-0001-6087-6643
Vivien J Bubb https://orcid.org/0000-0003-2763-7004

REFERENCES

1. Amarasinghe SL, Su S, Dong X, Zappia L, Ritchie ME, Gouil Q. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol* 2020;**21**:30
2. Vasieva O, Cetiner S, Savage A, Schumann GG, Bubb VJ, Quinn JP. Potential impact of primate-specific SVA retrotransposons during the evolution of human cognitive function. *Trends Evol Biol* 2017;**6**:DOI: 10.4081/eb.2017.6514
3. Kim K, Bang S, Yoo D, Kim H, Suzuki S. De novo emergence and potential function of human-specific tandem repeats in brain-related loci. *Hum Genet* 2019;**138**:661–72
4. Sulovari A, Li R, Audano PA, Porubsky D, Vollger MR, Logsdon GA, Human Genome Structural Variation C, Warren WC, Pollen AA, Chaisson MJP, Eichler EE. Human-specific tandem repeat expansion and differential gene expression during primate evolution. *Proc Natl Acad Sci U S A* 2019;**116**:23243–53
5. McGeary J. The DRD4 exon 3 VNTR polymorphism and addiction-related phenotypes: a review. *Pharmacol Biochem Behav* 2009;**93**:222–9
6. Kolla NJ, Vinette SA. Monoamine oxidase a in antisocial personality disorder and borderline personality disorder. *Curr Behav Neurosci Rep* 2017;**4**:41–8
7. Xiang C, Liu S, Fan Y, Wang X, Jia Y, Li L, Cong S, Han F. Single nucleotide polymorphisms, variable number tandem repeats and allele influence on serotonergic enzyme modulators for aggressive and suicidal behaviors: a review. *Pharmacol Biochem Behav* 2019;**180**:74–82
8. Botticelli L, Micioni Di Bonaventura E, Del Bello F, Giorgioni G, Piergentili A, Romano A, Quaglia W, Cifani C, Micioni Di Bonaventura MV. Underlying susceptibility to eating disorders and drug abuse: genetic and pharmacological aspects of dopamine D4 receptors. *Nutrients* 2020;**12**:2288
9. Blauw HM, van Rheenen W, Koppers M, Van Damme P, Waibel S, Lemmens R, van Vught PWJ, Meyer T, Schulte C, Gasser T, Cuppen E, Pasterkamp RJ, Robberecht W, Ludolph AC. Veldink JH, van den berg LH. NIPA1 polyalanine repeat expansions are associated with amyotrophic lateral sclerosis. *Hum Mol Genet* 2012;**21**:2497–502
10. Iacoangeli A, Al Khleifat A, Jones AR, Sproviero W, Shatunov A, Opie-Martin S, Alzheimer's Disease Neuroimaging I, Morrison KE, Shaw PJ, Shaw CE, Fogh I, Dobson RJ, Newhouse SJ, Al-Chalabi A. C9orf72 intermediate expansions of 24-30 repeats are associated with ALS. *Acta Neuropathol Commun* 2019;**7**:115
11. Tazelaar GHP, Boeynaems S, De Decker M, van Vugt JJFA, Kool L, Goedee HS, McLaughlin RL, Sproviero W, Iacoangeli A, Moisse M, Jacquemyn M, Daelemans D, Dekker AM, van der Spek RA, Westeneng H-J, Kenna KP, Assialioui A, Da Silva N, Povedano M, Pardina JSM, Hardiman O, Salachas F, Millecamps S, Vourc'h P, Corcia P, Couratier P, Morrison KE, Shaw PJ, Shaw CE, Pasterkamp RJ, Landers JE, Van Den Bosch L, Robberecht W, Al-Chalabi A, van den Berg LH, Van Damme P, Veldink JH, van Es MA, Project Min EALSSC. ATXN1 repeat expansions confer risk for amyotrophic lateral sclerosis and contribute to TDP-43 mislocalization. *Brain Commun* 2020;**2**:fcaa064
12. De Roeck A, Duchateau L, Van Dongen J, Cacace R, Bjerke M, Van den Bossche T, Cras P, Vandenberghe R, De Deyn PP, Engelborghs S, Van Broeckhoven C, Sleegers K, Consortium B. An intronic VNTR affects splicing of ABCA7 and increases risk of Alzheimer's disease. *Acta Neuropathol* 2018;**135**:827–37
13. De Roeck A, Van Broeckhoven C, Sleegers K. The role of ABCA7 in Alzheimer's disease: evidence from genomics, transcriptomics and methylomics. *Acta Neuropathol* 2019;**138**:201–20
14. Lam D, Ancelin M-L, Ritchie K, Freak-Poli R, Saffery R, Ryan J. Genotype-dependent associations between serotonin transporter gene (SLC6A4) DNA methylation and late-life depression. *BMC Psychiatry* 2018;**18**:1–10
15. Pacheco A, Berger R, Freedman R, Law AJ. A VNTR regulates miR-137 expression through novel alternative splicing and contributes to risk for schizophrenia. *Sci Rep* 2019;**9**:11793
16. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer

M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowki J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ, Szustakowki J. Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860–921

17. Bakhtiari M, Shleizer-Burko S, Gymrek M, Bansal V, Bafna V. Targeted genotyping of variable number tandem repeats with adVNTR. *Genome Res* 2018;**28**:1709–19
18. Castillo-Lizardo M, Henneke G, Viguera E. Replication slippage of the thermophilic DNA polymerases B and D from the euryarchaeota pyrococcus abyssi. *Front Microbiol* 2014;**5**:403
19. Audano PA, Sulovari A, Graves-Lindsay TA, Cantsilieris S, Sorensen M, Welch AE, Dougherty ML, Nelson BJ, Shah A, Dutcher SK, Warren WC, Magrini V, McGrath SD, Li YI, Wilson RK, Eichler EE. Characterizing the major structural variant alleles of the human genome. *Cell* 2019;**176**:663–75.e19
20. Linthorst J, Meert W, Hestand MS, Korlach J, Vermeesch JR, Reinders MJT, Holstege H. Extreme enrichment of VNTR-associated polymorphicity in human subtelomeres: genes with most VNTRs are predominantly expressed in the brain. *Transl Psychiatry* 2020;**10**:369
21. Nachman MW. Variation in recombination rate across the genome: evidence and implications. *Curr Opin Genet Dev* 2002;**12**:657–63
22. Pratto F, Brick K, Khil P, Smagulova F, Petukhova GV, Camerini-Otero RD. DNA recombination. Recombination initiation maps of individual human genomes. *Science* 2014;**346**:1256442–42
23. Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu Y-H, Kuhl DPA, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang F, Eussen BE, van Ommen G-JB, Blonden LAJ, Riggins GJ, Chastain JL, Kunst CB, Galjaard H, Thomas Caskey C, Nelson DL, Oostra BA, Warren ST. Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 1991;**65**:905–14
24. Kremer E, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren S, Schlessinger D, Sutherland G, Richards R. Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science* 1991;**252**:1711
25. Harley HG, Brook JD, Rundle SA, Crow S, Reardon W, Buckler AJ, Harper PS, Housman DE, Shaw DJ. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* 1992;**355**:545–46
26. Buxton J, Shelbourne P, Davies J, Jones C, Tongeren TV, Aslanidis C, Jong Pd Jansen G, Anvret M, Riley B, Williamson R, Johnson K. Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature* 1992;**355**:547–48
27. Aslanidis C, Jansen G, Amemiya C, Shutler G, Mahadevan M, Tsilfidis C, Chen C, Alleman J, Wormskamp NGM, Vooijs M, Buxton J, Johnson K, Smeets HJM, Lennon GG, Carrano AV, Korneluk RG, Wieringa B, Jong PJd. Cloning of the essential myotonic dystrophy region and mapping of the putative defect. *Nature* 1992;**355**:548–51
28. Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, Neville C, Narang M, Barcelo J, O'Hoy K, Leblond S, Earle-MacDonald J, de Jong PJ, Wieringa B, Korneluk RG. Myotonic dystrophy mutation: an unstable CTG repeat in the 3′ untranslated region of the gene. *Science* 1992;**255**:1253–5
29. Fu YH, Pizzuti A, Fenwick RG, Jr., King J, Rajnarayan S, Dunne PW, Dubel J, Nasser GA, Ashizawa T, de Jong P, Wieringa B, Korneluk R, Perryman MB, Epstein HF, Caskey CT. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* 1992;**255**:1256–8.
30. MacDonald ME, Ambrose CM, Duyao MP, Myers RH, Lin C, Srinidhi L, Barnes G, Taylor SA, James M, Groot N, MacFarlane H, Jenkins B, Anderson MA, Wexler NS, Gusella JF, Bates GP, Baxendale S, Hummerich H, Kirby S, North M, Youngman S, Mott R, Zehetner G, Sedlacek Z, Poustka A, Frischauf A-M, Lehrach H, Buckler AJ, Church D, Doucette-Stamm L, O'Donovan MC, Riba-Ramirez L, Shah M, Stanton VP, Strobel SA, Draths KM, Wales JL, Dervan P, Housman DE, Altherr M, Shiang R, Thompson L, Fielder T, Wasmuth JJ, Tagle D, Valdes J, Elmer L, Allard M, Castilla L, Swaroop M, Blanchard K, Collins FS, Snell R, Holloway T, Gillespie K, Datson N, Shaw D, Harper PS. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 1993;**72**:971–83
31. Orr HT, Chung MY, Banfi S, Kwiatkowski TJ, Jr., Servadio A, Beaudet AL, McCall AE, Duvick LA, Ranum LP, Zoghbi HY. Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat Genet* 1993;**4**:221–6
32. DeJesus-Hernandez M, Mackenzie Ian R, Boeve Bradley F, Boxer Adam L, Baker M, Rutherford Nicola J, Nicholson Alexandra M, Finch NiCole A, Flynn H, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung G-YR, Karydas A, Seeley WW, Josephs KA, Coppola G, Geschwind DH, Wszolek ZK, Feldman H, Knopman DS, Petersen RC, Miller BL, Dickson DW, Boylan KB, Graff-Radford NR, Rademakers R. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-Linked FTD and ALS. *Neuron* 2011;**72**:245–56
33. Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, Kalimo H, Paetau A, Abramzon Y, Remes AM, Kaganovich A, Scholz SW, Duckworth J, Ding J, Harmer DW, Hernandez DG, Johnson JO, Mok K, Ryten M, Trabzuni D, Guerreiro RJ, Orrell RW, Neal J, Murray A, Pearson J, Jansen IE, Sondervan D, Seelaar H, Blake D, Young K, Halliwell N, Callister JB, Toulson G, Richardson A, Gerhard A, Snowden J, Mann D, Neary D, Nalls MA, Peuralinna T, Jansson L, Isoviita VM, Kaivorinne AL, Hölttä-Vuori M, Ikonen E, Sulkava R, Benatar M, Wuu J, Chiò A, Restagno G, Borghero G, Sabatelli M, Heckerman D, Rogaeva E, Zinman L, Rothstein JD, Sendtner M, Drepper C, Eichler EE, Alkan C, Abdullaev Z, Pack SD, Dutra A, Pak E, Hardy J, Singleton A, Williams NM, Heutink P, Pickering-Brown S, Morris HR, Tienari PJ, Traynor BJ. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 2011;**72**:257–68
34. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes The Huntington's disease collaborative research group. *Cell* 1993;**72**:971–83
35. Ishiura H, Doi K, Mitsui J, Yoshimura J, Matsukawa MK, Fujiyama A, Toyoshima Y, Kakita A, Takahashi H, Suzuki Y, Sugano S, Qu W, Ichikawa K, Yurino H, Higasa K, Shibata S, Mitsue A, Tanaka M, Ichikawa Y, Takahashi Y, Date H, Matsukawa T, Kanda J, Nakamoto FK, Higashihara M, Abe K, Koike R, Sasagawa M, Kuroha Y, Hasegawa N, Kanesawa N, Kondo T, Hitomi T, Tada M, Takano H, Saito Y, Sanpei K, Onodera O, Nishizawa M, Nakamura M, Yasuda T, Sakiyama Y, Otsuka M, Ueki A, Kaida KI, Shimizu J, Hanajima R, Hayashi T, Terao Y, Inomata-Terada S, Hamada M, Shirota Y, Kubota A, Ugawa Y, Koh K, Takiyama Y, Ohsawa-Yoshida N, Ishiura S, Yamasaki R, Tamaoka A, Akiyama H, Otsuki T, Sano A, Ikeda A, Goto J, Morishita S, Tsuji S. Expansions of intronic TTTCA and TTTTA repeats in benign adult familial myoclonic epilepsy. *Nat Genet* 2018;**50**:581–90
36. Cortese A, Simone R, Sullivan R, Vandrovcova J, Tariq H, Yau WY, Humphrey J, Jaunmuktane Z, Sivakumar P, Polke J, Ilyas M, Tribollet E, Tomaselli PJ, Devigili G, Callegari I, Versino M, Salpietro V,

Efthymiou S, Kaski D, Wood NW, Andrade NS, Buglo E, Rebelo A, Rossor AM, Bronstein A, Fratta P, Marques WJ, Züchner S, Reilly MM, Houlden H. Biallelic expansion of an intronic repeat in RFC1 is a common cause of late-onset ataxia. *Nat Genet* 2019;**51**:649–58

37. Joensuu T, Kuronen M, Alakurtti K, Tegelberg S, Hakala P, Aalto A, Huopaniemi L, Aula N, Michellucci R, Eriksson K, Lehesjoki AE. Cystatin B: mutation detection, alternative splicing and expression in progressive myclonus epilepsy of Unverricht-Lundborg type (EPM1) patients. *Eur J Hum Genet* 2007;**15**:185–93
38. Tian Y, Wang J-L, Huang W, Zeng S, Jiao B, Liu Z, Chen Z, Li Y, Wang Y, Min H-X, Wang X-J, You Y, Zhang R-X, Chen X-Y, Yi F, Zhou Y-F, Long H-Y, Zhou C-J, Hou X, Wang J-P, Xie B, Liang F, Yang Z-Y, Sun Q-Y, Allen EG, Shafik AM, Kong HE, Guo J-F, Yan X-X, Hu Z-M, Xia K, Jiang H, Xu H-W, Duan R-H, Jin P, Tang B-S, Shen L. Expansion of human-specific GGC repeat in neuronal intranuclear inclusion Disease-Related disorders. *Am J Hum Genet* 2019;**105**:166–76
39. Marogianni C, Rikos D, Provatas A, Dadouli K, Ntellas P, Tsitsi P, Patrinos G, Dardiotis E, Hadjigeorgiou G, Xiromerisiou G. The role of C9orf72 in neurodegenerative disorders: a systematic review, an updated meta-analysis, and the creation of an online database. *Neurobiol Aging* 2019;**84**:238.e25–38.e34
40. Šerý O, Paclt I, Drtílková I, Theiner P, Kopečková M, Zvolský P, Balcar VJ. A 40-bp VNTR polymorphism in the 3′-untranslated region of DAT1/SLC6A3 is associated with ADHD but not with alcoholism. *Behav Brain Funct* 2015;**11**:21–21
41. Grünblatt E, Werling AM, Roth A, Romanos M, Walitza S. Association study and a systematic Meta-analysis of the VNTR polymorphism in the 3′-UTR of dopamine transporter gene and attention-deficit hyperactivity disorder. *J Neural Transm* 2019;**126**:517–29
42. Course MM, Gudsnuk K, Smukowski SN, Winston K, Desai N, Ross JP, Sulovari A, Bourassa CV, Spiegelman D, Couthouis J, Yu CE, Tsuang DW, Jayadev S, Kay MA, Gitler AD, Dupre N, Eichler EE, Dion PA, Rouleau GA, Valdmanis PN. Evolution of a human-specific tandem repeat associated with ALS. *Am J Hum Genet* 2020;**107**:445–60
43. MacKenzie A, Quinn J. A serotonin transporter gene intron 2 polymorphic region, correlated with affective disorders, has allele-dependent differential enhancer-like properties in the mouse embryo. *Proc Natl Acad Sci U S A* 1999;**96**:15251–55
44. Philibert RA, Wernett P, Plume J, Packer H, Brody GH, Beach SRH. Gene environment interactions with a novel variable monoamine oxidase a transcriptional enhancer are associated with antisocial personality disorder. *Biol Psychol* 2011;**87**:366–71
45. Tunbridge EM, Narajos M, Harrison CH, Beresford C, Cipriani A, Harrison PJ. Which dopamine polymorphisms are functional? Systematic review and meta-analysis of COMT, DAT, DBH, DDC, DRD1-5, MAOA, MAOB, TH, VMAT1, and VMAT2. *Biol Psychiatry* 2019;**86**:608–20
46. Bragg DC, Mangkalaphiban K, Vaine CA, Kulkarni NJ, Shin D, Yadav R, Dhakal J, Ton M-L, Cheng A, Russo CT, Ang M, Acuña P, Go C, Franceour TN, Multhaupt-Buell T, Ito N, Müller U, Hendriks WT, Breakefield XO, Sharma N, Ozelius LJ. Disease onset in X-linked dystonia-parkinsonism correlates with expansion of a hexameric repeat within an SVA retrotransposon in TAF1. *Proc Natl Acad Sci U S A* 2017;**114**:E11020–E28
47. Ord T, Puurand T, Ord D, Annilo T, Mols M, Remm M, Ord T. A human-specific VNTR in theTRIB3promoter causes gene expression variation between individuals. *PLoS Genet* 2020;**16**:20
48. Hu HY, He L, Fominykh K, Yan Z, Guo S, Zhang X, Taylor MS, Tang L, Li J, Liu J, Wang W, Yu H, Khaitovich P. Evolution of the human-specific microRNA miR-941. *Nat Commun* 2012;**3**:1145–45
49. Liu X, He J, Gao FB, Gitler AD, Fan D. The epidemiology and genetics of amyotrophic lateral sclerosis in China. *Brain Res* 2018;**1693**:121–26
50. MacKenzie A, Quinn JP. Post-genomic approaches to exploring neuropeptide gene mis-expression in disease. *Neuropeptides* 2004;**38**:1–15
51. Warburton A, Breen G, Bubb VJ, Quinn JP. A GWAS SNP for schizophrenia is linked to the internal MIR137 promoter and supports differential allele-specific expression. *Schizophr Bull* 2016;**42**:1003–08
52. Ali FR, Vasiliou SA, Haddley K, Paredes UM, Roberts JC, Miyajima F, Klenova E, Bubb VJ, Quinn JP. Combinatorial interaction between two human serotonin transporter gene variable number tandem repeats and their regulation by CTCF. *J Neurochem* 2010;**112**:296–306
53. Manca M, Pessoa V, Lopez AI, Harrison PT, Miyajima F, Sharp H, Pickles A, Hill J, Murgatroyd C, Bubb VJ, Quinn JP. The regulation of monoamine oxidase a gene expression by distinct variable number tandem repeats. *J Mol Neurosci* 2018;**64**:459–70
54. Manca M, Pessoa V, Myers P, Pickles A, Hill J, Sharp H, Murgatroyd C, Bubb VJ, Quinn JP. Distinct chromatin structures at the monoamine oxidase-A promoter correlate with allele-specific expression in SH-SY5Y cells. *Genes Brain Behav* 2019;**18**:e12483
55. Contente A, Dittmer A, Koch MC, Roth J, Dobbelstein M. A polymorphic microsatellite that mediates induction of PIG3 by p53. *Nat Genet* 2002;**30**:315–20
56. Martin P, Makepeace K, Hill SA, Hood DW, Moxon ER. Microsatellite instability regulates transcription factor binding and gene expression. *Proc Natl Acad Sci U S A* 2005;**102**:3800–4
57. Zukic B, Radmilovic M, Stojiljkovic M, Tosic N, Pourfarzad F, Dokmanovic L, Janic D, Colovic N, Philipsen S, Patrinos GP, Pavlovic S. Functional analysis of the role of the TPMT gene promoter VNTR polymorphism in TPMT gene transcription. *Pharmacogenomics* 2010;**11**:547–57
58. Vasiliou SA, Ali FR, Haddley K, Cardoso MC, Bubb VJ, Quinn JP. The SLC6A4 VNTR genotype determines transcription factor binding and epigenetic variation of this gene in response to cocaine in vitro. *Addict Biol* 2012;**17**:156–70
59. Kanno K, Ishiura S. Differential effects of the HESR/HEY transcription factor family on dopamine transporter reporter gene expression via variable number of tandem repeats. *J Neurosci Res* 2011;**89**:562–75
60. Burgoyne RD, Morgan A. Cysteine string protein (CSP) and its role in preventing neurodegeneration. *Semin Cell Dev Biol* 2015;**40**:153–59
61. Nosková L, Stránecký V, Hartmannová H, Přistoupilová A, Barešová V, Ivánek R, Hůlková H, Jahnová H, van der Zee J, Staropoli JF, Sims KB, Tyynelä J, Van Broeckhoven C, Nijssen PCG, Mole SE, Elleder M, Kmoch S. Mutations in DNAJC5, encoding cysteine-string protein alpha, cause autosomal-dominant adult-onset neuronal ceroid lipofuscinosis. *Am J Hum Genet* 2011;**89**:241–52
62. Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH. Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J Cell Biol* 1995;**128**:995–1002
63. Gendron TF, Bieniek KF, Zhang YJ, Jansen-West K, Ash PE, Caulfield T, Daughrity L, Dunmore JH, Castanedes-Casey M, Chew J, Cosio DM, van Blitterswijk M, Lee WC, Rademakers R, Boylan KB, Dickson DW, Petrucelli L. Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol* 2013;**126**:829–44
64. Zu T, Liu Y, Bañez-Coronel M, Reid T, Pletnikova O, Lewis J, Miller TM, Harms MB, Falchook AE, Subramony SH, Ostrow LW, Rothstein JD, Troncoso JC, Ranum LPW. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc Natl Acad Sci U S A* 2013;**110**:E4968–E77
65. Xi Z, Zhang M, Bruni AC, Maletta RG, Colao R, Fratta P, Polke JM, Sweeney MG, Mudanohwo E, Nacmias B, Sorbi S, Tartaglia MC, Rainero I, Rubino E, Pinessi L, Galimberti D, Surace EI, McGoldrick P, McKeever P, Moreno D, Sato C, Liang Y, Keith J, Zinman L, Robertson J, Rogaeva E. The C9orf72 repeat expansion itself is methylated in ALS and FTLD patients. *Acta Neuropathol* 2015;**129**:715–27
66. Bauer PO. Methylation of C9orf72 expansion reduces RNA foci formation and dipeptide-repeat proteins expression in cells. *Neurosci Lett* 2016;**612**:204–09
67. Nichol K, Pearson CE. CpG methylation modifies the genetic stability of cloned repeat sequences. *Genome Res* 2002;**12**:1246–56
68. Ziegler C, Domschke K. Epigenetic signature of MAOA and MAOB genes in mental disorders. *J Neural Transm* 2018;**125**:1581–88

69. Levens D. You don't muck with MYC. *Genes Cancer* 2010;**1**:547–54
70. Brooks TA, Hurley LH. Targeting MYC expression through G-quadruplexes. *Genes Cancer* 2010;**1**:641–49
71. Zamiri B, Mirceta M, Bomsztyk K, Macgregor RB, Jr., Pearson CE. Quadruplex formation by both G-rich and C-rich DNA strands of the C9orf72 (GGGGCC)8•(GGCCCC)8 repeat: effect of CpG methylation. *Nucleic Acids Res* 2015;**43**:10055–64
72. Savage AL, Bubb VJ, Breen G, Quinn JP. Characterisation of the potential function of SVA retrotransposons to modulate gene expression patterns. *BMC Evol Biol* 2013;**13**:101–01
73. Hall A, Moore AK, Hernandez DG, Billingsley KJ, Bubb VJ, Quinn JP. Nabec North American brain expression consortium. A SINE-VNTR-Alu in the LRIG2 promoter is associated with gene expression at the locus. *Int J Mol Sci* 2020;**21**:8486
74. Ewing AD, Smits N, Sanchez-Luque FJ, Faivre J, Brennan PM, Richardson SR, Cheetham SW, Faulkner GJ. Nanopore sequencing enables comprehensive transposable element epigenomic profiling. *Mol Cell* 2020;**80**:915–28e5